Vol. 62, No. 5 September, 1955

Psychological Review

THEODORE M. NEWCOMB, Editor
UNIVERSITY OF MICHIGAN
Lorraine Bouthilet, Managing Editor

CONTENTS

Principles of Performance.....EDWARD C. TOLMAN 315

A Note on Second-Order Conditioning—
and Secondary Reinforcement.....GREGORY RAZRAN 327

Psychology from a Biologist's
Point of View.....C. JUDSON HERRICK 333

The Stimulus Conditions Which Follow
Learned Responses.....CHARLES C. PERKINS, JR. 341

Consciousness, Behavior, and
Personality.....JOSEPH NUTTIN 349

The Difference Between *Q* and *R*.....JACK BLOCK 356

Some Problems in the Theory
of Vigilance.....JAMES DEESE 359

Statistical Theory of Distributional
Phenomena in Learning.....W. K. ESTES 369

Relationships Between Shame and Guilt
in the Socializing Process.....DAVID P. AUSUBEL 378

Notes Toward a Motor Theory of
Visual Egocentric Localization.....JAN H. BRUELL
AND GEORGE W. ALBEE 391

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

SOLOMON E. ASCH
ROBERT R. BLAKE
STUART W. COOK
CLYDE H. COOMBS
LEON FESTINGER
W. R. GARNER
JAMES J. GIBSON
D. O. HEBB
HARRY HELSON
E. R. HILGARD
CARL I. HOVLAND
E. LOWELL KELLY
DAVID KRECH
ROBERT W. LEEPER
ROBERT B. MACLEOD
DAVID C. MCCLELLAND
GEORGE A. MILLER
GARDNER MURPHY
OSCAR OESER
CARL PFAFFMANN
CARROLL C. PRATT
DAVID SHAKOW
RICHARD L. SOLOMON
ELIOT STELLAR
S. S. STEVENS
ERIC TRIST
EDWARD L. WALKER
ROBERT W. WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STREETS, LANCASTER, PA.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

$6.50 volume $1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright © 1955 by the American Psychological Association, Inc.

VOL. 62, NO. 5 SEPTEMBER, 1955

THE PSYCHOLOGICAL REVIEW

PRINCIPLES OF PERFORMANCE [1]

EDWARD C. TOLMAN

University of California

Apparently it behooves a so-called "cognitive theorist" such as myself to explain how his assumed intervening cognitive and motivational variables ever issue into actual behavior. I must, it seems, provide some *rules* or *principles* of *performance* according to which these variables—namely, discriminations, expectancies, drives, incentive values—add to or multiply with one another (or whatever the appropriate combining operations may be) to produce behavior. Otherwise, as Guthrie (**5**) has suggested, I seem to leave my animal (whether rat or human being) merely "buried in thought" (though why this should be considered so shocking in the case of a human being I have never quite been able to fathom).

Let me begin by summarizing briefly my basic assumptions. First, I assume four distinctive types of *independent variables:* (*a*) present stimulus-units; (*b*) previous presentations of stimulus-unit–response–stimulus-unit sequences, abbreviated as past (*S-R-S*)'s; (*c*) present drive disturbances and currently presented incentive objects; and (*d*) genetic or early training differences between individual subjects (as parameters affecting the results from the other three types of independent variable).

[1] Presented, with some modification, as presidential address to the Division of General Psychology of the American Psychological Association, Sept. 7, 1953.

Secondly, I assume three types of *intervening variables* and *intervening-variable systems:* (*a*) discriminations within a discrimination system; (*b*) beliefs and expectancies within a belief-expectancy system; and (*c*) drives and incentive-values within a drive-incentive-value system. The three intervening variable systems are inferred sets of dispositions (in the sense of propensities of the organism to perform in certain prescribable ways under specified test conditions). And I assume that the momentary activated discriminations, beliefs, expectations, drives, and incentive values are determined by the momentary values of the presented independent variables in the contexts of these three innate and acquired dispositional systems.[2] Let me elaborate.

1. *The Discrimination System.* This is to be specified by (*a*) *categorizing* or *nominal dimensions,* that is, types of general object-classification, of which the given organism has become capable (e.g., most human beings have acquired

[2] I wish to express my great indebtedness to Professor Leo Postman for his unstinting and crucial help in the development of these concepts concerning the independent and the intervening variables and variable systems. The performance principles which I shall propose later in this article, though in part Professor Postman's, are at present largely mine—at least insofar as the pictorial language is concerned. Postman does not find this language helpful.

such categories as colors, pitches, smells, chairs, tables, trees, other human beings, etc., etc.), and by (*b*) the *intensitive dimensions* along which the given organism has, in addition, become capable of making intensity discriminations (e.g., specific hues, heights of pitches, intensity of odors, comfortableness of chairs, sturdiness of tables, shadiness of trees, likeableness of human beings, etc.). Such categorizing and intensitive dimensions are uncovered in standard situations in which the organism is presented with selected arrays of stimulus-units and which it is forced (or tricked) into categorizing and/or discriminating intensitively. In other words, the discrimination system is to be discovered by the methods of a comprehensive and up-to-date, as well as a mathematically sophisticated, psychophysics. Knowing this discrimination system, we can thereafter state the specific categorizations and intensity discriminations which are likely to be activated when such and such stimulus-units are actually presented on any given occasion.

2. *The Belief-Expectancy System.* This is to be specified by the sum of the organism's expectancy dispositions. For example, a subject may show that upon the sounding of a given tone, S_1, he is prepared, provided he waits, R_1, in a given place, for the coming of a shock, S_2. Or again, he may show that when presented by the choice point of a maze, S_1, he is prepared for a resultant food, S_2, provided he turns, R_1, to the right. Or he may show that upon the presentation of a first syllable, S_1, on a memory drum he is prepared, provided he continues to look at the drum, R_2, for the coming of a second syllable, S_2. Any one such set in the animal, which (to distinguish it from the objective occurrences) may be symbolized as an $(\dot{S}_1\text{-}\dot{R}_1\text{-}\dot{S}_1)$, I would designate as an activated expectation. The continued possession by the organism of such a set I would call an expectancy disposition. An "expectancy disposition" is thus a small acquired disposition of greater or less certainty and permanence, which will tend to lead to an activated expectation whenever an instance of the given sign stimulus-unit is presented. Such an expectancy disposition will also tend to be activated by other stimulus-units which are near the sign stimulus on some common dimension of generalization. Thus, there may be groups or families of expectancy dispositions in the sense that an array of them related by different signs may all have attached to them similar expectancy dispositions as to the type of significate to be reached, if a certain general type of means-manipulatory response be carried out. Such a group or envelope of related expectancy dispositions I shall call a "belief disposition." An individual's expectancy system is constituted by the sum of all his expectancy and belief dispositions.[3]

3. *The Drive-Incentive-Value System.* This system is to be conceived as a collection of dispositions to approach and/or consume specific types of positive incentive objects or to escape from and/or avoid specific types of negative incentive objects. These approach (consummation) and avoidance (and/or escape) dispositions depend upon the interaction of two distinguishable independent variables: namely, (*a*) types and amounts of drive-disturbance, and (*b*) types of stimulus-units presented as incentive objects.

a. Drive-disturbances would be: hours of specific food-deprivation, hours of sex-deprivation in males, time-position of the female in the oestrus cycle, per-

[3] I would like to emphasize that my notions about expectancies have been clarified by the recent discussion by MacCorquodale and Meehl. I wish here to acknowledge my very great indebtedness to them (13).

centage decrement in bodily weight, hours of water-deprivation, amounts and prior time of saline injections, intensities of just-preceding electric shocks, hours of social deprivation, specific raisings and lowerings of bodily health, special maintenance-schedule with respect to aesthetic objects or with respect to exercise, etc., etc.

b. Incentive objects would be: specific foods, liquids, sex objects, groups of individuals of the same or different species, barriers, and distances to be overcome, enemies, friends, present electric shocks, poems, paintings, symphonies, open windows to look through in the case of monkeys (see Harlow [6]), etc.

The character and strength of an activated approach (consummation) and avoidance (escape) disposition will result from the combination of the particular drive-disturbance and the particular incentive object presented on the given occasion. The sum total of drive-disturbance and incentive object relations obtaining for a given individual make up his drive-incentive-value system.

Assuming now that the above brief discussion of discrimination dispositions, expectancy dispositions, and drive-incentive-value dispositions has succeeded in presenting the basic theoretical position, I wish to pass on to the main topic of this paper. How on particular occasions—given the independent variables of presented sign stimulus-units, previously presented (*S-R-S*)'s, present drive-disturbances and present positive or negative incentive objects—do the to-be-inferred resulting then-and-there activated discriminations, expectancies, drives, and incentive values interact to produce actual performance? What, in short, are the Rules or Principles of Performance which must be added to the system?

PERFORMANCES AND THE BEHAVIOR SPACE

I shall illustrate these principles or rules in large part by means of what I shall call behavior-space diagrams or pictures. This notion of a behavior space stems, of course, from Kurt Lewin's concept of the life space (12). It differs from the latter, however, *first*, in that the structures and processes which Lewin put inside his "psychological person" *as part* of the life space have been, in the present analysis, taken care of by the three dispositional systems of discriminations, belief-expectancies, and drives and incentive values, together with the momentary processes activated within these systems.

Secondly, my behavior space is more frankly a merely pictorial way of representing certain relationships and less a would-be true mathematics or logic.

Thirdly, I differ perhaps from Lewin in that what I am trying to illustrate or explain by the behavior space are not concrete responses but certain generalized *performances*. A "performance," as I here use the term, is not something that can be observed on a single occasion but something that can be specified only from observing two or more concrete behaviors. For example, Tolman, Ritchie, and Kalish in one of their experiments on spatial learning (25) used a single-unit **T** maze with two vertically opposite, alternatively usable, starting stems. And they discovered that only by observing the responses of the animals from both stems could they say whether a rat's acquired behavior—what I am here calling his "performance"—was that of "taking right-hand turns" or that of "going to the lighted side." But, whichever it was, such an "always going to the right or "always going towards the lighted side," would

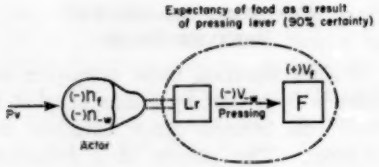

FIG. 1. Behavior-space: rat pressing lever for food in a Skinner box.

be an example of what is here meant by a performance.[4]

Fourthly, it is to be emphasized that a behavior-space diagram will be neither a picture of the actual geographical environment nor a picture of the supposed phenomenological experience of the organism. It is rather a merely pictorial way of trying to represent the interactions between the inferred activated dispositions and their consequent determination of the to-be-predicted performance. For example, my diagram for the case of the rat pressing the lever in a Skinner box (17) at a certain rate is shown in Fig. 1.

This diagram represents the assumption that the rat discriminates himself as "actor" in the presence of the discriminated lever, *Lr*. The pair of dotted lines connecting actor to *Lr* symbolizes "immediate presence of the lever." And the diagram represents further the fact that this discriminated lever, *Lr*, has become the sign term in the activated expectancy. The whole expectancy is indicated by the contents of the dash-dot ellipse. This expectancy has acquired, it is assumed, a certainty of about 90 per cent and is to the effect that if the actor responds by the "performance" of pressing, he will get to food, *F*. Also, the expected food is depicted as having a positive valence v_f, indicated by the plus sign. And the actor is shown as having an activated drive or need-push *for* food, indicated by the minus sign, labeled n_f, inside him. Further, the length of the dashed arrow, labeled "pressing," indicates the degree of effort or work which, according to the expectancy, will be involved in the pressing. And the actor is shown as having a drive or need-push *against* work indicated by the second minus sign, labeled n_{-w} within the actor. Finally, the arrow at the left, labeled *Pv*, is a vector representing the magnitude and direction of the resulting to-be-predicted performance of pressing.

Let us consider further for a moment need-pushes and valences. I conceive a need-push as, so to speak, that portion of a drive which, under the concrete stimulus conditions of the moment, gets into the behavior space. The magnitude of a need-push is, in short, assumed to be determined not only by the magnitude of the actual activated drive but also by the goodness of the discrimination of drive stimuli under the given experimental conditions.

Consider, for example, that type of latent-learning experiment which has been called by Thistlethwaite (19) the irrelevant-incentive experiment. In such experiments, during the initial training period an irrelevant goal object (say food) is presented in one alley of a T or Y maze, and water, a relevant goal object (since the animals are in this period run thirsty), is presented in the other alley. Then the rats are made hungry and it appears from the original experiment of Spence and Lippitt (18) and from various repeat experiments (see, for example, Kendler [11]) that rats often do *not* then tend to go to the food side when first made hungry. However, it has also appeared from an experiment by Thistlethwaite (20) that,

[4] For this notion of a performance as a behavioral disposition which can be identified only by observing at least two somewhat different test situations I am indebted primarily to Professor Benbow F. Ritchie. (Private communication.) But see also a quite recent statement of what seems to be essentially the same position by Campbell (2).

if the animals have had some previous training in discriminating their drive stimuli, they are much more likely to show some initial tendency to go at once to the food side. These two contrasting results would be interpreted, by me, as meaning that, although the rats have been made hungry (as would be shown by their readiness to approach or consume food in the standard situation), this hunger drive does not tend to be discriminated (i.e., to produce a corresponding need-push in the behavior space corresponding to the maze) unless there has been previous practice in drive discrimination in a mazelike situation. Further, some recent findings by Heyer and O'Kelly (7) indicate that tissue dehydration due to subcutaneous saline injection will not cause appropriate performance in a maze to get to water, *unless* this dehydration has been experienced over a considerable preceding period of time. I would interpret this also as lack of thirst need-push in the behavior space (except in the actual presence of water itself) unless there has been plenty of preceding time to experience and discriminate the special drive stimuli resulting from the saline-produced dehydration.

Return now to the Skinner box. We may probably assume that in it, because of preliminary training, the hunger drive stimuli and the nonwork drive stimuli have become completely discriminated. We may assume in short that the magnitude of the food need-push corresponds directly to the magnitude of the actual food hunger drive. And finally we may probably also assume that the magnitude of the nonwork need-push corresponds directly to the actual nonwork drive.

What now about the valences? Valences are conceived as bearing nearly the same relation to incentive-values that need-pushes bear to drives. That is, the amount of valence which the food pellets will have in the behavior-space diagram corresponding to the Skinner box will depend not only upon the incentive value of such pellets (as measured in some standard situation) but also upon the discriminability of the pellets in the concrete stimulus context of a Skinner box. And this discriminability may, hypothetically at least, be less than in the standard situation. Hence the valence of the food, in the behavior space corresponding to the Skinner box, could be less (though I doubt it) than the basic incentive value of such pellets, as measured in a standard situation. Similarly, the negative valence of the work involved in pressing might, in the behavior space corresponding to the Skinner box, be less than the actual negative incentive value of such work. But this also seems doubtful.

Look again at Fig. 1. My final argument is that as a result of (i) the need-push for food, (ii) the positive valence of the expected food, (iii) the need-push against work, and (iv) the negative valence of the expected work (this expected work is indicated by the length of the dashed arrow), there results the performance vector *Pv*. It is this vector which specifies the direction and magnitude of the to-be-predicted actual performance. The greater the valence of the expected food, the greater the food need-push, and the greater the expectancy that the food will result, the greater the magnitude of the performance vector toward actually pressing the lever. On the other hand, the greater the need-push against work and the greater the negative valence of the expected work and the stronger the expectation of this work, the less the performance vector toward pressing the lever.

For the sake, now, of the critic who dislikes such a pictorial mumbo jumbo, it may be pointed out that these relations probably can be equally well ex-

pressed by an algebraic equation (Equation 1).

$$Pv = f_x\ (n_f,\ v_f,\ \exp_f) - f_y\ (n_{-w},\ v_{-w},\ \exp_w) \quad [1]$$

Here Pv stands for the magnitude of the performance vector (i.e., the tendency to press the lever). And this magnitude is shown as increasing with the values of n_f, v_f, and $\exp_f$, combined according to some function, f_x, and as being decreased by the values of n_{-w}, v_{-w}, and $\exp_w$, combined according to some function, f_y. The exact forms of these two functions, f_x and f_y, I have not attempted to specify. It may be that these functions both have the form of simple multiplication. But in any event, it is assumed for both of these functions that the component variables all operate in a positive direction such that the increase in any one of these component variables tends to increase the value of the function.

Still further, the critic will note that the first half of the right-hand side of this equation is rather similar to Hull's equation for reaction potential, or ${}_sE_R$ (9), (which is presented here as Equation 2).

$$\text{(Hull)}\ {}_sE_R = f\ (D \times V_1 \times K \times {}_sH_R) \quad [2]$$

Hull's D, or drive, is closely similar to my n_f, or hunger need-push. His V_1, or stimulus intensity dynamism, is, however, a term I do not have. For me the effect of any *intensity* or a presented stimulus-unit such as the stimulus-pattern resulting from the lever would already have been evaluated by an analysis of the discriminatory dispositional system and of the effects in it of intense or nonintense stimulus-units upon the discriminations of Lr and of F. Hull's K, or incentive motivation, corresponds to my valence, v_f. And his ${}_sH_R$, or habit, is an S-R connection whereas my expectancy is conceived as an activated (S_1-R_1-S_2) expectation. Further, however, it is to be noted that Hull's equation does not contain the second part of the right-hand side of my equation. It may be recalled, however, that he does introduce, separately, the construct of conditioned inhibition ${}_sI_R$ which, if I interpret him correctly, is somewhat related to what I am here saying about w or expected work. ${}_sI_R$ he apparently subtracts from ${}_sE_R$ just as I subtract $f_y\ (n_{-w},\ v_{-w},\ \exp_w)$. It must be pointed out, however, that for Hull inhibition is somewhat different from my concept of expected work and need against work. But in any event, both he and I assume that the "expected work" or the "conditioned inhibition" gets established through repetitions of the activity involved.

Turn now to escape behavior. I would feel that, in the usual case, escape is not as has been suggested by the Hullians primarily a getting-to a positive valence attached to the reduction of pain but is rather simply a getting-away from the negative valence of the shock itself. Such an interpretation is represented by Fig. 2. In Fig. 2 the jumping is represented as a mere leading away from the grid. And the animal will merely jump in any direction easily available.

Turn next to avoidance behavior. Avoidance behavior has always been something of a problem for S-R psychologists. This point was made some years ago by Hilgard and Marquis (8) and has recently been re-emphasized and subjected to a further analysis by

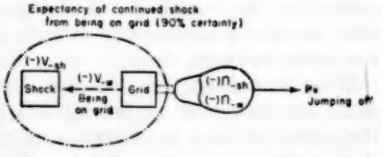

FIG. 2. Behavior-space: rat escaping from charged grid.

Ritchie (16). In avoidance behavior the performance vector results, according to my language, from the expectation that going to the grid is going to lead to shock. And this performance vector pushes or pulls the animal away from the grid even when he is not on it. The negative need-push against shock will lead, not as in the escape case merely to getting off the grid after the shock has come, but rather to the performance of avoiding the grid altogether. See Fig. 3. In such a case, pain never occurs and hence there is no pain to be reduced. Mowrer (14), of course, is the one among the stimulus-response-oriented psychologists who has most clearly seen this. And he has solved it by postulating a conditioned anxiety to the whole grid box, the reduction of which constitutes positive reinforcement. Now while I would accept something like Mowrer's interpretation for many instances, I do not believe that anxiety is always or necessarily involved. The skilled human subject who has learned to lift his finger in time to avoid a minor shock is not usually, I believe, suffused with even a minimum of anxiety. He responds so as to avoid the shock, quite unemotionally and purely on the basis of his expectancy, his need-push, and the negative valence of the shock. He expects shock in a certain place at a certain time, provided he goes to or stays in that place, and his performance is one of avoiding that place before that time, in a quite relaxed, nonanxious fashion.

Finally, however, it is to be noted that although my *diagrams* for escape and avoidance are different from that for approach, my algebraic equation for either escape or avoidance would be of the same form as that for approach (Equation 3).

$$Pv = f_x\,(n_{-sh},\, w_{-sh},\, \exp_{sh}) - f_y\,(n_{-w},\, v_{-w},\, \exp_w) \quad [3]$$

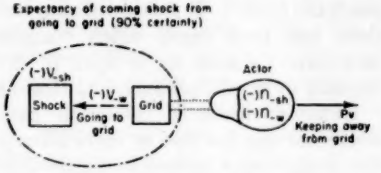

FIG. 3. Behavior-space: rat avoiding grid from which shock is expected.

The greater the need-push n_{-sh} against shock, the greater the negative valence v_{-sh} of the particular shock and the more certain the expectation $\exp_{sh}$ (or immediate discrimination of shock), the stronger will be the performance vector leading to behavior away from the sign of pain (in the case of avoidance) or leading to the leaving of the pain-object (in the case of escape). And the greater the need against work n_{-w}, the greater the negative valence v_{-w} of the particular work involved in the escape or the avoidance and the greater the expectancy of this work, the smaller will be the Pv toward escape or avoidance.

It is to be noted further, however, that, in the avoidance case, w, or the expected work involved between sign and significate, can be large and lead to relatively weak avoidance of the sign as a result of a number of relatively different physical reasons. The w may be waiting a long time for the shock. Or it may involve doing some special further act such as pressing a treadle for the shock to appear. Or, to take a quite different example, where the significate is not pain from shock but pain from a rat-hungry cat, w might involve the breaking down of a tough barrier by the rat between it and the cat. In all three cases the avoidance response of the rat will be relatively weaker than it would have been if these intervening w's had been small.

But up to now you may well feel that my formula and Hull's are not significantly different. And, if you do, so

much the better. It would indicate that there has been really much common agreement between us in spite of differences of words, phrases, and slogans.

However, when it comes to cases where not one but two or more alternative performance tendencies are involved and result in some quite new performance—such, for example, as taking a compromise path between two original ones, taking a shortcut, taking a roundabout route, VTEing, or the like—then depicting the situation by means of the behavior space becomes (to me) more helpful and suggestive. (And, of course, not to renege on my basic position, my own "private style of thinking," to quote Kendler [11], is such that I cannot easily think about the interaction of a single sign, expectancy, need-push, and valence without first seeing these in some sort of spatial terms.) But turn now to the new problem of combining two expectancies. I shall introduce here a new concept, that of the angular dispersion between expectancies.

This concept of angles is tricky and I know I am not wholly clear on it. It is one that Lewin also bogged down on. Nevertheless let me present, as an illustration, one type of compound performance which I would explain with the help of angles, namely, VTEing (vicarious trial and error). The explanation will be closely related to, but in some ways different from, that which Clark Crannell and I evolved some years ago (21, 22) in connection with what I then called the "schematic sowbug." Actually, anyone who read the earlier papers may have already recognized that the actor as I have presented him here is no more than a modern, and I hope a less complicated, version of my old friend, the schematic sowbug.

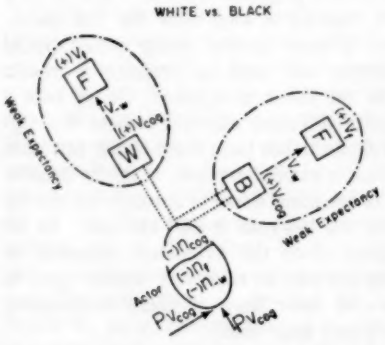

FIG. 4a. Behavior-space, early in discrimination learning. White card versus black card. Much VTE.

There are two sets of empirical findings as to VTEing to be emphasized: (*a*) Early in learning, just before (or just as) the rats begin to react to the crucial sign dimension, they tend to do more VTEing when there is a large discriminable difference between the two cards than when there is a small one. They VTE more for a white card vs. a black card than for a white card vs. a light gray card (21). (*b*) Later, however, when they have had much preceding experience in discriminating along the white-gray-black dimension, they tend to VTE more when the discriminable difference between the two cards is small (24). They VTE more for white vs. light gray cards than for white vs. black cards.

Consider first the behavior-space pictures early in learning when the rats VTE more for the white-black than for the white-gray. (See Fig. 4a and Fig. 4b.)

In Fig. 4a the actor (the sowbug) is in the white-black situation and in Fig. 4b in the white-gray one. And I have assumed that white vs. black corresponds to a 90° angle relative to the actor's (sowbug's) nose, whereas the white vs. gray corresponds to a smaller angle. I have also assumed that in this early part of learning there operates, in addition to the hunger need-push n_f, a more purely cognitive sort of need-push

to examine new stimuli, n_{cog}. The need-push corresponding to this cognitive-examination drive I have indicated by a minus sign in the fore part of the actor or bug. I have also assumed that both the stimulus cards have for a while, as new stimuli, positive valences for this cognitive-examining drive. That is, there will be performance vectors pushing the actor to "look at" each stimulus card. These vectors are those indicated at the rear of the bug. In addition, I have indicated, by putting in F (for food) with a plus valence on it on both sides, that at this early stage of learning there have developed expectancies that going toward either card may lead to food. And for the diagram to be complete there should also be another pair of shortish vectors pushing the bug, as a result of these expectancies, to jump towards each card. I have omitted these latter, jumping vectors, so as not to complicate the picture. The vectors shown are "looking towards and examining" vectors.

I shall assume further (and here I am assuming very much what Glanzer [3, 4] has proposed in two recent articles in which he was concerned with alternation behavior in a symmetrical maze) that in the course of looking at one card the valence of this card as "something to look at and examine" tends to become temporarily discharged or satiated. This then weakens, for the moment, the vector for looking at this first card. The actor then turns to look at the other card. The vector toward this latter card, however, then becomes satiated in its turn; whereas the vector for looking at the first card becomes stronger again. So the rat looks back and forth, or VTE's, until satiation for both cards, as far as looking or examining is concerned, has become complete. In short, the rat VTE's until jumping takes over.

To explain further why there is more VTEing for white vs. black (Fig. 4a) than for white vs. gray (Fig. 4b) I shall also assume that since in the case of white vs. black the two cognitive-examination vectors are at right angles, the reduction of one of these vectors through satiation has no effect on the other. The rat gets temporarily satiated for white but he still wants to look at black, and vice versa. In the case of the white vs. gray, on the other hand, the two examining vectors are at a narrower, acute, angle which means that the shortening of one vector through satiation projects upon and tends to shorten the other. Looking at white reduces temporarily not only the vector to look at white but also the one to look at gray, and vice versa. Hence early in learning or just before learning there will be less VTEing for white vs. gray than for white vs. black.

Consider now the situation later in learning. The original cognitive-examination drive is assumed now to have become wholly satiated. The vectors corresponding to it will have disappeared. But expectations have begun to build

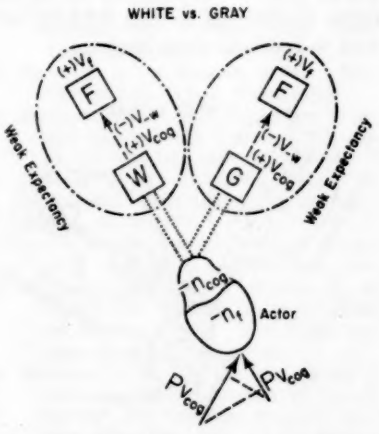

FIG. 4b. Behavior-space, early in discrimination learning. White card versus gray card. Little VTE.

up to the effect that, in the one case, white leads to food and black to no food, and, in the other, that white leads to food and gray to no food. These situations are represented in Figs. 5a and 5b.

The vectors are no longer ones for looking, but ones for jumping to white (as sign of food) and for not-jumping to black (as sign of nonfood). In the white-black case (Fig. 5a), because the two vectors are at right angles the negative vector "away from the black" does not affect the positive one "toward the white," and vice versa. In the white-gray case (Fig. 5b), on the contrary, since the two arrows are at an acute angle, the negative vector "away from the gray" does tend to project upon and reduce the positive vector "toward the white"; and the positive one "toward the white" projects upon and decreases the negative one "away from the gray." And this means in the white-gray case that until the expectancies become very certain and the vectors very long, the reduction of the positive vector by the negative one and vice versa will be sufficient to keep the rat in unstable equilibrium so that VTEing will tend to continue much longer.

Now, for the sake of the nonvisually minded critic, let me here suggest that perhaps a more respectable and conventional way of talking would be to say merely that early in learning there is more generalization of satiation of the cognitive examining drive from white to gray than from white to black; whereas, later in learning, there is more generalization of acquired approach and avoidance tendencies from white to gray than from white to black. More generalization of satiation for "looking at" early in learning means *less* VTEing between white and gray, whereas more generalization of approach and avoidance tendencies later in learning means *more* VTEing between white and gray. Actually, however, I suspect that such a more "respectable" way of talking is really just as confused and as full of hidden assumptions as is my pictorial way. In any event, the pictorial terms suggest to me (and I emphasize *to me*) a greater number of further experiments to try out than do the purely verbal ones.

The topic of this paper was "Principles of Performance." And by a per-

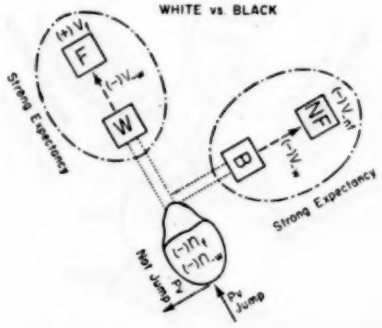

FIG. 5a. Behavior-space, later in discrimination learning. White card versus black card. Little VTE.

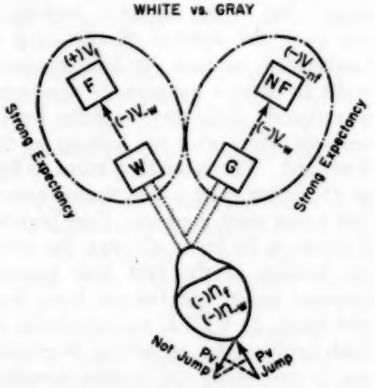

FIG. 5b. Behavior-space, later in discrimination learning. White card versus gray card. Much VTE.

formance was meant, as has been said, a generalized way of behaving to be discovered and specified only by observing the responses in at least two and perhaps more concrete test situations. Further, a performance when thus discovered is identifiable only in general terms, such as approach, consummation, escape, avoidance, vicarious trial-and-error examination, or the like. The motor skill characteristics involved in such a performance are not included in its definition. Only in a very general way may a performance be further specified as a "looking at," a "jumping to," a "running down," a "pressing," an "eating," or a "drinking." The point of view being argued for is that performances and not movements or skills are all that we ought now, at the present time, to try to explain or to predict in detail. Under certain controlled conditions we can use movements as pointer readings for performances but we cannot as yet predict movements independently of the performances within which they occur. As Brunswik has pointed out (1), all kinds of different movements and skills may vicariously substitute for one another and produce one and the same achievement or performance. Movements and skills in his terms are mere mediational, peripheral phenomena. It is the "distal," not the "proximal," characters of *S*'s and *R*'s to which I am suggesting that we devote ourselves for the time being. This is my interest and my bias.

One last word. Most of the really interesting problems of psychology—such as the distortion effects of motivation on perception (and I still believe there are some) or again those facts which have to do with what I have called the widening or narrowing of cognitive maps (23)—will eventually find their place, I believe, under the heading of complications in the principles of performance. Or, to use the language of this paper, they will turn out to be describable in terms of interesting complications and interactions between the performance vectors, as these complications and interactions appear in particular behavior-space setups.

REFERENCES

1. BRUNSWIK, E. Organismic achievement and environmental probability. *Psychol. Rev.*, 1943, **50**, 255–272.
2. CAMPBELL, D. T. Operational delineation of "what is learned" via the transposition experiment. *Psychol. Rev.*, 1954, **61**, 167–174.
3. GLANZER, M. The role of stimulus satiation in spontaneous alternation. *J. exp. Psychol.*, 1953, **45**, 387–393.
4. GLANZER, M. Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.*, 1953, **60**, 257–268.
5. GUTHRIE, E. R. *The psychology of learning* (Rev. Ed.). New York: Harper, 1952.
6. HARLOW, H. F. Motivation as a factor in the acquisition of new responses. In *Current theory and research in motivation, a symposium*. Lincoln: Univer. of Nebraska Press, 1953, pp. 24–49.
7. HEYER, A. W., JR., & O'KELLY, L. I. Studies in motivation and retention. *Comp. Psychol. Monogr.*, 1951, **20**, Serial No. 106.
8. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: Appleton-Century-Crofts, 1940.
9. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
10. KENDLER, H. H. An investigation of latent learning in a T-maze. *J. comp. physiol. Psychol.*, 1947, **40**, 265–270.
11. KENDLER, H. H. "What is learned?" A theoretical blind alley. *Psychol. Rev.*, 1952, **59**, 269–277.
12. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.
13. MACCORQUODALE, K., & MEEHL, P. E. Section 2 in Estes, W. K. *et al. Modern learning theory*. New York: Appleton-Century-Crofts, 1954.
14. MOWRER, O. H. On the dual nature of learning—a reinterpretation of "conditioning" and "problem-solving." *Harvard Educ. Rev.*, 1947, **17**, 102–148.
15. MURRAY, H. A. *Explorations in personality*. New York: Oxford Univer. Press, 1938.

16. RITCHIE, B. F. Can reinforcement theory account for avoidance? *Psychol. Rev.*, 1951, **58**, 382–386.
17. SKINNER, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton-Century-Crofts, 1938.
18. SPENCE, K. W., & LIPPITT, R. An experimental test of the sign-gestalt theory of trial-and-error learning. *J. exp. Psychol.*, 1946, **36**, 491–502.
19. THISTLETHWAITE, D. A critical review of latent learning and related experiments. *Psychol. Bull.*, 1951, **48**, 97–129.
20. THISTLETHWAITE, D. Conditions of irrelevant-incentive learning. *J. comp. physiol. Psychol.*, 1952, **45**, 517–525.
21. TOLMAN, E. C. Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.*, 1939, **46**, 318–336.
22. TOLMAN, E. C. Discrimination vs. learning and the schematic sowbug. *Psychol. Rev.*, 1941, **48**, 367–382.
23. TOLMAN, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, **55**, 189–208.
24. TOLMAN, E. C., & MINIUM, E. VTE in rats: overlearning and difficulty of discrimination. *J. comp. Psychol.*, 1943, **36**, 91–98.
25. TOLMAN, E. C., RITCHIE, B. F., & KALISH, D. Studies in spatial learning: II. Place learning versus response learning. *J. exp. Psychol.*, 1946, **36**, 221–229.

(Received November 22, 1954)

Psychological Review
Vol. 62, No. 5, 1955

A NOTE ON SECOND-ORDER CONDITIONING—AND SECONDARY REINFORCEMENT

GREGORY RAZRAN

Queens College

In *Principles of Behavior* (12, ch. 7), Hull set up second-order conditioning as a sort of fulcrum of his much used construct of secondary reinforcement. And, except for Skinner (30, p. 245), other CR learning theorists seem to regard it, too, as quite a general CR manifestation (19, 22, 28, 31). Indeed, one gets the impression that most of these theorists subsume second-order conditioning under secondary reinforcement in the sense that an S-R or R that is able to reinforce another S-R or R must *ipso facto* be able to condition (in a classical manner) the other S-R or R to itself (19, pp. 438, 457; 22, pp. 430–433; 28; 31, pp. 697–699). Frolov's experiment (7), reporting that the sight of a black square paired with a conditioned sound of a metronome came to produce by itself conditioned salivation, is widely quoted in textbooks and in integrative summaries and reviews of the topic.

Yet the fact remains that, as directly observed in the laboratory, second-order conditioning is a rather uncommon and unstable phenomenon. Of the 347 Russian reports on classical conditioning of dogs and monkeys that the writer has managed to see only eight (4, 6, 7, 9, 14, 16, 17, 33) [1] offer accounts of any degree of successful second-order conditioning. Much more commonly—in 63 separate investigations—nonreinforced pairings of conditioned and nonconditioned stimuli are reported as resulting in conditioned inhibition; that is, the second-order to-be-conditioned stimuli, instead of evoking a CR come to interfere with it when they are applied simultaneously or shortly before the first-order conditioned stimuli. (In an even larger number of cases, as might be expected, the Russian protocols show that the pairing produces no effect at all; or, as Pavlovians usually put it, "the animal's nervous system was too excitable" [or insufficiently inhibitable —one would seem to have it either way in that system] "to develop the conditioned inhibition.") True, a greater frequency of second-order conditioning may be found in Russian experiments with lower animals (hermit crabs 18; fish 8); with school children (10, 20, 23, 29—more often with Ivanov-Smolensky's instrumental than with Krasnogorsky's classical technique); and with adult human subjects (1, 13, 15, 21). But even here, results are at times sporadic and the finding itself poses some special questions. Moreover, no bona fide second-order conditioning has ever come out of any American laboratory,[2] and the writer has never been able to demonstrate it in any of his series of studies with "misinformed" subjects in whom most other behavioral Pavlovian findings have been corroborated.

[1] Reference to the original reports of the experiments by Froloff and by Lindberg, cited respectively in Pavlov's *Conditioned Reflexes* (7, p. 27) and in Pavlov's *Pavlovian Wednesdays* (16, pp. 196, 201, 224, 241), could not be found in any Russian bibliographical periodical. The writer assumes that such reports have never been published.

[2] The experiments by Culler and Finch (5) and by Brogden and Culler (2) are a somewhat special case, inasmuch as the higher-order to-be-conditioned stimuli were combined with a thoracic shock as a "motivator" or reinforcer (pedal shock was the unconditioned stimulus).

If second-order conditioning is thus not a general but a very special CR manifestation, two hypotheses would seem to be suggested by the studies in which it reportedly has manifested itself. One—which the writer intends to reject—regards second-order conditioning as merely an artifact of reactivated CR *stimulus generalization* or, as Woodworth and Schlosberg term it, CR *stimulus nondifferentiation*—a more descriptive term than the writer's CR stimulus *pseudo-generalization*. It might, for instance, be argued that, while nonreinforced applications of nonconditioned stimuli by themselves improve, through contrasts, CR stimulus differentiation, nonreinforced combined applications of nonconditioned and conditioned stimuli (method of second-order conditioning) on the contrary reactivate, through confusion, CR stimulus nondifferentiation. Hence: second-order stimuli may well evoke CR's for which they originally were ineffective. Reactivated CR stimulus nondifferentiation is quite frequently reported by Russian laboratories when subjects are administered some drug or undergo surgery or succumb to illness or are merely confronted with a difficult task of differentiation (also, after a lapse of time) —and presumed second-order conditioning might conceivably be another example of it.

To be sure, such an assertion could be readily checked by series of CR nondifferentiation tests before and after second-order training sessions. But, unfortunately, no account of such tests could be found in the Russian CR reports, and their finding of more second-order conditioning in lower animals (whose capacity for differentiation and for learning in general is admittedly underdeveloped) than in dogs and monkeys would seem to indicate further an "artifact-nondifferentiation" view. Yet this view is strongly opposed by two separate sources of evidence. (*a*) In three otherwise well-controlled Russian studies—two with dogs (**4**, **16**) and one with monkeys (**17**)—second-order conditioning persisted in its original form even after its first-order conditioning was "transswitched," i.e., changed from food to defense CR's, and vice versa. (*b*) All second-order CR's in dogs and monkeys were best formed—whenever they were formed at all—when their first-order CR's were well established and hence at a minimum of stimulus nondifferentiation. Consequently, the hypothesis of second-order conditioning as an artifact and a case of stimulus nondifferentiation might better be rejected as an over-all interpretation of the phenomenon, even though it might still be retained as a plausible account of its alleged manifestation in fish and invertebrates.

The second hypothesis, taking its cue from Russian CR studies of school children, would maintain that second-order conditioning, unlike first-order conditioning, requires some sort of organismic cognition to be established. And here the writer would like to present direct evidence from one of his own experiments. In this experiment, food was used as the unconditioned stimulus, the sound of a metronome and the flashing of a green light were the first-order and second-order to-be conditioned stimuli, and salivation was the unconditioned reaction. The 15 adult human subjects in the experiment were divided equally into a "misinformed" group and two "informed" ones: one "informed without suggestion" and the other "informed with suggestion." The misinformed group was told in the first-order conditioning that "the purpose of the experiment is to study the effects of auditory fatigue upon digestion" and in the second-order conditioning that they were "tested to find out which affects digestion more, auditory or visual

stimuli." On the other hand, the informed-without-suggestion group was instructed about Pavlov's work with dogs and told that "the purpose of the experiment is to find out whether human subjects condition similarly," while the informed-with-suggestion group was told that human subjects *do* condition like Pavlov's dogs and that "the purpose of the experiment is to gather more data on the subject."

The results—scores in terms of *t* values between conditioned and control salivation—are presented in Table 1 and are quite clear-cut. *The informed groups, particularly the informed-with-suggestion group, manifested a considerable amount of second-order conditioning, whereas the misinformed group, whose first-order conditioning was by all tokens more regular and greater than that of the informed groups did not.*

The hypothesis that second-order conditioning needs cognition to be adequately effected is not for the present concerned with the nature of the cognition needed. The distinction between cognition and noncognition is so far a purely global-operational one: the differences obtained between subjects who know and who do not know that they are being conditioned. Specific conceptualizations of processes involved in the cognition might better be held in abeyance until after specific experiments have been devised to test them. Moreover, the hypothesis does not of course deny such cognition to animals. All it maintains is that in typical experiments of classical conditioning, animal subjects manage only occasionally to exercise adequate cognition while as a rule they fall short of it simply because they do not cognize what their human experimenters want them to cognize—or do. (In instrumental or operant conditioning, in the writer's estimation, animal subjects "catch on" more readily since there, as often stated, the to-be-conditioned or to-be-reinforced S-R or R produces, and not merely accompanies, the unconditioned or reinforcing S-R or R.)

TABLE 1

SECOND-ORDER SALIVARY CR'S OF 15 ADULT HUMAN SUBJECTS AS DEPENDENT UPON THEIR COGNITION OF THE CONDITIONING

(Scores are *t* values ($df = 35$) between conditioned and control salivations, and are in all cases results of three 50-min. training sessions.)

Stimuli Tested	Groups Tested		
	Misinformed	Informed Without Suggestion	Informed With Suggestion
1. Metronome (first-order stimulus)	24.5	2.9	15.4
2. Green light (second-order stimulus)	.6	1.9	3.6

Elsewhere (25) the writer has postulated—upon the basis of an array of experimental evidence—a central cognitive factor in stimulus generalization; and here, in second-order conditioning, such a factor seems to him even more likely to be operative. And in both cases, it needs to be repeated, the argument rests upon the same two convergent considerations: (*a*) the differences between "knowing" and "unknowing" human subjects (partly also the "introspections" [sic!] of the subjects) with respect to the CR characteristic in question as found by the writer; and (*b*) the irregularities in the characteristic as evident in the large number of studies with animal subjects in the Russian laboratories. Basic irregularities in a phenomenon like conditioning, the data of which come from several hundred well-controlled experiments of a uniform design, cannot, the writer posits, be resolved by merely negating the phenomenon or ignoring the irregularities. Rather, they call for supplementary as-

sumptions, which is what is proposed here.

Now, if second-order conditioning is difficult to obtain and maintain in animals, and it seems to need cognition to be set up in human beings, what about secondary reinforcement per se, conceived independently as the capacity of a conditioned stimulus to reinforce more or less permanently another reaction? In the writer's opinion, the answer is essentially the same. True, the laboratory evidence for secondary reinforcement in animals—while never too critically appraised—seems, to date, to be more consistent than that for second-order conditioning. Yet, as an isolated force, secondary reinforcement is, too, as any mere inspection of the data of its several experiments[3] reveals, very impermanent, very extinguishable, and quite ancillary and immediately dependent upon primary reinforcement. Thus there is here, too, a far cry from the results obtained (several experiments) to the phenomena (the vast area of man's cultural behavior) which are purported to be explained—and lots of room is left for supplementary accounts.

Twenty years ago, in a critical analysis of the entire CR area in phylogeny and ontogeny, the writer suggested what he then called a "reservoir theory" of relative CR permanence and "unextinction"—viz.: ". . . the emergence somewhere in the evolutionary scale of an organismic pattern, perhaps the cortical representation of the vocal or proprioceptive parts of the conditioning [unconditioned] and conditioned stimuli, that acts as a reservoir of associational 'energy' and makes the strength and existence of a connection or C-R independent of the 'energy of' or 'reinforcement by' the original stimulus that caused the C-R or connection to be formed" (**24**, p. 108). And more recently, in studying the "interiorization of conditioned attitudes," he found that "to attain some permanence, to stay conditioned, conditioned attitudes require some post-conditioning verbalization which seems to act as some sort of consolidator of the conditioned change" (**26**, p. 3). Consequently, it would seem reasonable to assume that secondary reinforcement, too, may fall within the category that needs verbalization and/or cognition to become a part of man's permanent behavior repertoire.[4]

On the other hand, the writer would contend that conditioning in which the unconditioned reaction or the primary reinforcer is a dominant biological activity—that is, primary or first-order classical conditioning or primary instrumental or operant conditioning—need not invoke cognition for its initiation and maintenance (even if cognition may well affect its operation). The basis for such a contention is plain. Classical conditioning is quite effective in a variety of instances in which the assumption of cognition would be patently gratuitous. Consider, for instance, the conditioning of decorticated and curarized animals (**11**), of paramaecia (**32**), of coelenterates (**34**), and, above all, the large volume of successful conditioning of all types of visceral reflexes,

[3] The "reinforcement" of secondary reinforcement must obviously go beyond the mere occurrence of the conditioned reaction, thus excluding a number of experiments that have been bracketed with it. A guinea pig running away from a cage in which he had been shocked, or a dog retrieving an object that had been associated with food, or even a college student reporting a food sensation to a conditioned metronome, are cases not of secondary reinforcement but of primary conditioning

[4] It need also be mentioned that nonclassical reinforcement learning (or conditioning) admittedly takes place in settings in which, compared with classical conditioning, cognition is more likely to be available and more likely to be utilized (skeletal vs. autonomic action; choice, or at least more choice of, response).

including viscerovisceral ones, that the Russians have reported in recent years (cf. **3**, **27**). And, while there is no evidence for "primitive" primary conditioning of the instrumental-operant type, there would seem to be—at least for the present—no methodological advantage and no "ontological argument" for invoking cognition as the *sine qua non* for its coming into being.

But all this does not in any way negate the hypothesis of the cognitive nature of second-order classical conditioning and of secondary reinforcement. On the contrary, it corroborates a view that the writer has held for some time—namely, that the learning of anything, whether it be a movement or a motive or an act or an attitude, must depend directly on either some source of specific biological energy or some center of general cognitive direction (whatever the ultimate nature of the energy—or the correlated energy—of the latter). In learning, the writer is an unabashed dualist.

REFERENCES

1. BEKHTEREV, V. M., & DOBROTVORSKAYA, N. I. [The development of chained associated reflexes.] *Vopr. Izuch. Vospit. Lichn.*, 1927, **3–4**, 1–10.
2. BROGDEN, W. J., & CULLER, E. Experimental extinction of higher order reflexes. *Amer. J. Psychol.*, 1935, **47**, 663–669.
3. BYKOV, K. M. [*The cerebral cortex and the internal organs.*] Moscow: Medgiz, 1947.
4. CHEBOTAREVA, O. [Further materials in the physiology of conditioned reflexes.] Thesis, St. Petersburg, 1912.
5. CULLER, E., & FINCH, G. Higher order conditioning with constant motivation. *Amer. J. Psychol.*, 1934, **46**, 596–602.
6. DOROZDENKO, N. P. [The effector path of conditioned reflexes of the second order.] *Fiziol. Sh. SSSR.*, 1950, **36**, 519–523.
7. FROLOV, I. P. Quoted in I. P. Pavlov, *Conditioned reflexes.* New York: Oxford Univer. Press, 1927. Pp. 34–35.
8. FROLLOF, J. P. [Frolov, I. P.] Bedingte reflexe bei Fischen, II. Mitteilung. *Pflüg. Arch. Physiol.*, 1928, **220**, 339–349.
9. FURSIKOV, D. S. [The relation between the processes of excitation and inhibition.] *Trud. Fiziol. Lab. Pavlova*, 1924, **1**, 1–47.
10. GACKEL, L. [An investigation of conditioned inhibition in hysterical children.] *Zh. Nevropatol. Psikhiat.*, 1928, **21**, 682–690.
11. GIRDEN, E. Conditioned responses in curarized monkeys. *Amer. J. Psychol.*, 1947, **60**, 571–587.
12. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
13. ISRAELSON, Z. I. [Topical differentiation of associated motor reflexes to dermal and muscular stimuli.] *Izv. Imper. Voyenno-Med. Akad.*, 1914, **28**, 73–78.
14. KASHERININOVA, N. A. [Contributions to the study of salivary conditioned reflexes to tactile stimuli in the dog.] Thesis, St. Petersburg, 1908.
15. LENZ, A. [Conditioned reflexes of higher order and their study in mental patients.] *Russ. Fiziol. Zh.*, 1922, **5**, 93–.
16. LINDBERG, A. A. Quoted in I. P. Pavlov, [*Pavlovian Wednesdays.*] Moscow and Leningrad: Akademiya Nauk, 1949. Vol. 1: pp. 196, 201, 224, 240; Vol. 2: p. 229.
17. MALINOVSKY, O. V. [Conditioned reflexes of the second order in monkeys.] *Trud. Inst. Fiziol. Pavlova*, 1952, **1**, 205–212.
18. MIKHAILOFF, S. Expériences réflexologiques. Expériences nouvelles sur Pagurus striatus, Leander xiphias et treilleans. *Bull. Inst. Oceangr.*, 1923, No. 422, 1–16.
19. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 435–472.
20. NOVIKOVA, A. A. [Conditioned inhibition and conditioned reflexes of "higher order" in children.] *Med.-biol. Zh.*, 1929, **1**, 120–131.
21. OPARINA, N. [The relations of the processes of excitation and inhibition during the action of extra stimuli.] *Nov. Refleksol. Fiziol. Nervn. Sist.*, 1929, **3**, 196–209.
22. OSGOOD, C. E. *Method and theory in experimental psychology.* New York: Oxford Univer. Press, 1953.
23. PANFEROV. Quoted in N. I. Krasnogorsky, Bedingte und unbedingte Reflexe in Kindesalter und ihre Bedeutung für die

Klinik. *Ergebn. inner. Med. Kinderhk.*, 1931, **39**, 613–730.

24. RAZRAN, G. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, 1935, No. 191.
25. RAZRAN, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365.
26. RAZRAN, G. The interiorization of conditioned attitudes. Paper read at East. Psychol. Ass., New York, 1954.
27. RAZRAN, G. Conditioning and perception. *Psychol. Rev.*, 1955, **62**, 83–95.
28. SEWARD, J. P. Secondary reinforcement as tertiary motivation: a revision of Hull's revision. *Psychol. Rev.*, 1950, **57**, 362–374.
29. SHASTIN, N. R. [Unconditioned and conditioned reflexes in myxedema.] *Russ. Fiziol. Zh.*, 1930, **6**, 470–482.
30. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
31. SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 690–729.
32. TCHAKHOTINE, S. Réactions conditionnés par micro-puncture ultra-violette dans le comportement d'une cellule isolée (*Paramecium caudatum*). *Arch. Inst. Prophyl.*, 1938, **10**, 119–133.
33. ZELIONY, G. P. Espèce particulière de réflexe conditionnel. *Arch. Sci. Biol., St. Petersburg*, 1909, **15**, 437–453.
34. ZUBKOV, A. A., & POLIKARPOV, G. G. [Conditioned reflexes in coelenterates.] *Usp. Sovr. Biol.*, 1951, **32**, 301–302.

(Received October 8, 1954)

Psychological Review
Vol. 62, No. 5, 1955

PSYCHOLOGY FROM A BIOLOGIST'S POINT OF VIEW

C. JUDSON HERRICK

Grand Rapids, Michigan

The Basic Problem

A student of the philosophy of science might be tempted to say there are as many biologies as there are biologists, if account is taken only of the men whose exceptional insight and productiveness have guided the growth of the science. The same may be true of psychology, and perhaps we can do no better than accept Cattell's characteristic definition—"Psychology is what the psychologist is interested in *qua* psychologist." There are, however, some general principles about which biologists and psychologists are in substantial agreement. One of these principles is that no factual findings have scientific significance until they are fitted into the appropriate niche in the integrated system of knowledge.

The neurologist finds this orientation especially difficult because almost all experiences and activities of men and other animals involve nervous functions and his field has no boundaries. The human brain is the most important thing in the world, for, as Gibbs (1) expresses it, "Human history is a history of the brain activity of the human race" (p. 1505). This relationship ties neurology closely with psychology and also with psychiatry, sociology, and every other human interest. But when the neurologist tries to find out just where his findings tie in with psychology he is puzzled. There are so many psychologies that one wonders what it is all about.

The interested spectator who sits on the fence watching the game sees two opposing teams and, on the side lines, a goodly number of other psychologists who do not join either faction. In one team the partisans of traditional dualism contend for a sharp separation of the conscious, or "spiritual," activities from the unconscious, or "physical," thus splitting the world as we experience it into two universes, one of which has been characterized as "spiritual reality" or "ideational reality" and the other as "physical reality." Opposed to these radical spiritists are the mechanists, who insist that, since the search in both science and philosophy is for unifying principles, and since it has not been possible to explain how a nonphysical agency can act upon a physical structure so as to influence human conduct, we must search for physicalistic principles of sufficiently wide import to embrace all the known phenomena.

The more radical members of the second group, apparently accepting the traditional doctrine that anything "spiritual" is *ipso facto* nonphysical, take the easy way out and deny that conscious experience of any kind has scientific or operational significance. This despite the fact that the very denial is a conscious act. This exclusion of everything mentalistic from psychology is obviously a defense reaction against the primitive animistic mythologies which still survive in every human culture. But even though mind is called an epiphenomenon, it is nonetheless a phenomenon, a natural event, and a place in the system of nature must be found for every natural event.

The spiritists' quest for a psychology released from the limitations imposed by the laws of the physical world, and the objective psychologists' insistence that only observable physical processes

are significant for psychology, seem to be irreconcilable. The controversy points again to the fact that the basic problem of psychology is, as it always has been, the exact nature of the relationship between our knowledge of the objective world and the subjective experience of knowing and all other conscious acts.

My purpose here is to examine, from the standpoint of mechanistic biology, some of the diverse fields of inquiry which must be integrated before this basic problem can be formulated in accordance with physical, physiological, and psychological principles now generally accepted. The divisive tendencies of current scientific movements are retarding progress.

The key factor in the current controversy about the nonphysical nature of the human spirit really hinges on a definition: Just what do we mean by physical structure and what are its properties? We must define the physical before we can talk rationally about the nonphysical. A brief summary of a few principles of current physical science is prerequisite to further consideration of the meaning of the word "physical" in biological and psychological contexts.

PHYSICAL SCIENCE OF TODAY

The history of ideas about the nature of the physical world records three revolutionary periods. The Greek period, typified by Aristotle and Euclid, dominated formal thinking for two thousand years and is still influential. Beginning in the middle ages a second revolutionary period culminated with the Newtonian system of mechanics. The third period began with the twentieth century, as exemplified in relativity and quantum mechanics. It is still in its infancy and its findings are more revolutionary than any of the preceding.

The fundamental conceptions of physics are in flux. New methods have revealed new facts which require new principles, and some principles formerly regarded as axiomatic are now suspect. The absolutes of earlier times are now treated relativistically, with radical reorganization of the science of mechanics.

To put it briefly, natural science today regards our cosmos as a stupendous mechanism (physicalism) composed of innumerable subsidiary mechanisms, all bound together in accordance with lawfully ordered principles. We do not adequately understand any thing or any event until we know the mechanism that produces it and the principles in accordance with which it operates. A mechanism is defined by D'Arcy Thompson (7) as ". . . whatsoever checks or controls, and guides into determinate paths, the workings of energy" (p. 291). Accepting this definition, the science of mechanics deals with energy and the "whatsoever" that controls its workings.

The mechanism makes some specific kind of product and the nature of the product is the crucial issue in its organization. This product may be material arranged in a different way or place, or energy in a different pattern of manifestation. Or it may be matter transformed into energy or energy into matter, for these are known to be interconvertible in quantitatively measurable relations (Einstein's conversion equation, $E = Mc^2$). In view of this last point the distinction between matter and energy becomes rather fuzzy, and any manifestation of an energy change is a physical event.

The belief now current among physicists is that the various kinds of atoms are relatively stable and different patterns of energy. There are no different kinds of energy. The so-called thermal, electrical, and other energies are differ-

ent patterns of manifestation of one common measurable quantity which in our ignorance we call energy. Nature as a whole is process. There is nothing static about it anywhere.

In a natural mechanism the materials and energies used may come from a wide field and the product made is in turn delivered to the surroundings. It follows that the active structure, the machine itself, must be so defined as to include the entire field with which it has transactional relations. The mistaken popular notion that a machine is a passive inert structure operated by an outside agent is derived from the artificial machine with a human operator. It is true that a lathe in a machine shop has no value as mechanism without the operator who starts it, stops it, and controls its action. This means that, from the operational standpoint, the operator must be regarded as an integral part of the mechanism. In natural mechanisms it is more evident that the operator is inside the apparatus. The apparatus itself is the operator; that is to say, the machine operates itself. No natural mechanism needs a djinn or any external operator to run it or tell it what to do. This, we believe, is as true of a man as of a solar system or a volcano.

Classical physics as formulated in Newtonian mechanics deals with inert solid particles of matter which differ in mass and are pushed about by forces acting upon them. These forces are considered to be manifestations of energy, and the movements are measurable in arbitrary units of space and time. Twentieth century science, on the contrary, finds that matter and energy are different manifestations of the same unknown something and that in some domains of physics space and time in the objective world must be treated relativistically with reference to each other. In subatomic physics they cannot be measured in arbitrary units, with the observer as a fixed point of reference.

A physical mechanism is defined dynamically. It may make some particular product, repetitively, because it is so organized as to do this by virtue of this organization. But if the organization changes its pattern so as to deliver a different kind of product, a factor is introduced which may properly be called creative. Even a repetitive performance like that of some particular chemical reaction exhibits the property of transforming a pattern of material or energy into a different pattern, and this capacity is the source from which creativity, as here defined, is derived. Creation does not imply that something is made out of nothing. The scientific problem is to discover the laws in accordance with which these changes take place. Since the operating forces are manifestations of energy, it is evident that energy as such has creative efficiency. In other words, its activities are not stereotyped in rigidly predetermined patterns. These patterns are constantly changing by conversion as they interact with one another, and this capacity for change is responsible for cosmic and organic evolution and for the orderly processes of growth of living individuals. This perhaps is what Whitehead had in mind when he said, "creativity is ultimate"; that is to say, it is something which science cannot explain or analyze further.

This intrinsic creativity of physical processes is not a metaphysical speculation. It is an observed fact, and if it were not so our cosmos would now be in a state of homogeneous equilibrium at the lowest level of organization. In view of these dynamic and creative properties of the physical universe, and of our incomplete knowledge of the laws in accordance with which the successive changes take place, we can set

no limits to the kinds of products that natural mechanisms (physicalisms) can make. If now we wish to push further into the domain of psychophysics, account must be taken of this change in the climate of science which has come within the memory of men now living.

The modern mechanist is not a materialist in the classical sense. We grant that he deals only with physical structures and their operations. The operating body is always a mobile structure, which may be relatively stable as we see it in a mountain or a cat, or it may be a specific pattern of energy manifestation in which no particulate matter is recognizable, as in a ray of light. Usually it is a combination of both, as in a volcano or a thinking man. In current science by "mechanistic" we mean physicalistic, and these words now have meanings which are quite different from those of fifty years ago, although the latter are still current in both common speech and scientific literature.

Subsequent references to relativity and quantum mechanics do not carry the implication that the psychobiologist thinks that these principles, as now formulated in mathematical physics, are adequate to give a satisfactory explanation of the mind-body relationship. The most that can be claimed for them in this connection is that they may be steps toward the discovery of the still unknown principles that satisfy the requirements of the problem.

THE NATURAL HISTORY OF THE SPIRIT

My understanding of current scientific thought about some of the properties of the natural world has just been outlined. It is physicalistic throughout. Let us see now where the opposing psychologies mentioned at the beginning of this essay stand in the light of this evidence.

Many years ago I said (2) that introspection is as manifestly a biological method as is reflexology, and this principle needs re-emphasis today. Conscious experience influences conduct. It is therefore a factor of behavior. We may go further and say with Edwin G. Boring that a conscious experience is itself a behavior, a bodily act which may be observed introspectively just as we observe other kinds of behavior. Introspection, like the other methods of observation, is, of course, subject to the hazard of erroneous interpretation. And again it is timely to insist that the dualistic theories merely pile mystery upon mystery and only confuse the issues without explaining anything.

Because man is an animal and because his "spiritual" capacities are demonstrably vital processes, psychology is necessarily articulated with biology; but it does not follow from this that the laws of conscious experience are identical with those of the things of which we have experience. It is, in fact, clear that they are not. Each level of organization has its own distinctive properties, some of which cannot be reduced to those of lower levels (3).

The mind-body problem will never be solved by ignoring the troublesome factors, either those of spirit or of matter. The inquiry cannot be limited to either the conscious or the unconscious factors, because what we are looking for is the relation between the two. It does no good to try to evade the issue as a "pseudoproblem." Traditional materialism (the "crude" variety) and classical spiritism (or, more reputably, "idealism") both involve neglect of a vast wealth of human experience, including common sense and refined scientific knowledge. We cannot choose between materialism and spiritualism. We must have both.

The attempts so far made to find a

suitable formula of psychophysics have led into blind alleys. The chief difficulty arises from the fact that we have no common measure for objectively known dimensions and the relations observed in the process of knowing them. If ever a suitable formula for this relationship is discovered, it will probably be stated in some kind of relativistic terms, and it may dispense with numerical units altogether. Mathematics and symbolic logic are now developing methods of dealing with situations devoid of any quantitative factors, and this may point the way toward hitherto unexplored fields of inquiry in psychobiology.

The human brain controls many of the internal bodily activities and by far the larger part of overt behavior. Its structure is inconceivably complicated and much of the detail is still obscure. There is on record, however, a vast amount of knowledge about nervous structure and the laws of its operation. The products delivered by this mechanism are of many kinds, including growth of the structure itself, a variety of chemical reactions and associated changes in electrical potential, and the excitation of mass movements of the body. These have been thoroughly studied and described in physiological terms.

There is another unique property of some of these operations that cannot be studied by the objective methods of anatomy and physiology, because it is strictly private and can be recognized only subjectively by the person who is operating. Fortunately, we have a mechanism by which this personal awareness can be made public. Ideas can be translated into symbols—words and others—the utterance of which evokes similar symbols in other people. By this indirection subjective experience can be objectified and so communicated. This enables us to develop a legitimate science of introspective psychology. We do not know the mechanics of this process of symbolizing or the physical laws of its operation, but we have an interesting analogy in the electronic computing machines which perform similar operations with a quite different mechanism. Doubtless there are some common principles in these two kinds of operations.

There are no conscious factors, so far as we know, in most physiological operations; but we know as surely as we know anything in science that all mental acts are vital functions and we know where the organs are that perform them. We know also that the distinctively human types of higher mental functions are added to pre-existing physiological functions in the course of personal and evolutionary development.

Although we do not yet know how awareness emerges within the metabolism of brain tissue, we know that it does so and that it is as truly a physical process as is the transmission of light through apparently empty space. We do not know all about the mechanics of light either, but long ago Helios was banished from science because he could not explain anything. When a weight is lifted by muscular action the cause of the movement is not a disembodied contractility. It is contracting muscle. When a problem in mental arithmetic is solved the operating agent is not a nonphysical mind which activates an inert structural mechanism. It is a living brain engaged in thinking. There is no specific kind of mental energy. Mental work is bodily work, and it is the body that gets tired when we do it. Mind does not move matter. It is "minding" matter that does "mental" work just as it is contracting muscle that does "physical" work. A thought is not a product made by a mechanism, the way bile is

made by the liver. It is the operation of the mechanism, an act, not a material or any other kind of an object. *To split the function from the organ that performs it is scientifically as absurd as is the separation of the properties of any other thing from the object that manifests them.*

The claim of the parapsychologists that psi phenomena in general involve "extrasensory perception," presumably through nonphysical agencies, calls for a critical appraisement of current studies of the mechanics of perception. This inquiry brings out very clearly the fundamental difference between the relations of space and time factors in the acquisition of perceptual knowledge of the objective world and these relations in the higher rational use of this knowledge.

Every perception of a physical object or event and every consciously controlled movement of the body involves a transactional relation of some sort between the "spiritual" and the "physical." In all these activities the conscious and the unconscious factors are inextricably interwoven. The first can be observed only introspectively, the other again introspectively but only by an indirection, by the observer becoming aware of some objective (unconscious) physical events. The act of perception involves the conversion of "physical" processes into "mental" processes, and conversely a voluntary act requires the conversion of a conscious purpose into movement of the physical body. The nub of the mind-body problem lies in the nature of these conversions.

The significant factors of such conversions can be isolated by experimental analysis more readily in the study of perception than of any other psychological processes. We have biological evidence that the internal integrative and regulatory functions of an animal body must be sharply distinguished from the analytic functions of sensorimotor type which are concerned with the adjustment of the body to external things and events. The physical principles of integration are fundamentally different from those of the analytic functions, and some of the former can be described only in relativistic terms.

The most complicated integrative apparatus known is in the human brain, and the act of knowing is an integrative process which gives us two kinds of knowledge which Sellars (6) has called "perceptual knowledge" and "conceptual knowledge." The former gives us all the information we have about the spatial, temporal, and energic properties of the objective world. Newtonian mechanics was developed on the basis of this knowledge. This mechanics is of necessity framed in behavioral space and time. Conceptual knowledge comprises those higher rational and emotive functions which cannot be quantified in the numerical units of space, time, mass, and energy of Newtonian mechanics.

This distinction is fundamental. It is explained, in part at least, by the nature of the act of perception and of the apparatus employed. Its neglect accounts for the failure of the earlier attempts to find acceptable laws of psychophysics. This theme cannot be elaborated here. Enough has been done to show that this lead points to a promising field worthy of further investigation.

PARAPSYCHOLOGY

The intimate relationship between objective things and events and our knowledge of them is so commonplace that it is generally uncritically accepted. But the manifestations of hypnotism, telepathy, and the other so-called psi phenomena are so unusual, and in some instances so bizarre, that there is in-

sistent demand that they be either explained or explained away as spurious. The latter has been tried repeatedly without convincing proof.

These phenomena have been under investigation by competent scientists for more than seventy-five years. The workers in this field of "parapsychology" who are now most active are advocates of dualistic theories which have been given wide publicity, but the majority of other scientists question the significance of their findings. This skepticism is due in part to the fact that this field of inquiry was so long contaminated with fraud and delusion that the prejudice against it is still strong. The prejudice has not been mitigated by the fantastic claims made in some of the recent literature. Nevertheless, these are real phenomena which merit critical study by every available method.

It should be recognized and emphasized that the mysteries associated with psi phenomena are of the same order, and no different in principle, as those presented by every mental and intentionally controlled act whatsoever. If parapsychology is to win acceptance as legitimate natural science, it must be fitted into the same general frame of reference as the other branches of psychology and of natural science as a whole. "Parapsychology" is a misnomer, for psi phenomena are either psychological or they are not. If not, what are they?

In psi phenomena the unconscious factors clearly play the critical role. These can be studied only by the methods of objective science and these methods have so far proved inadequate. The same is true for many other unsolved scientific problems. So when Dr. Rhine writes (**5**, p. 300), "We know by this time that we cannot use a physical theory to explain psi," we freely grant that no satisfactory physical theory has yet been found; but that is far from proof that it cannot be found. In another passage (**4**, p. 62) he has insisted that the physics of tomorrow is irrelevant here; we must think in terms of present knowledge. But this is exactly what he does not do. What he does do is to base his argument on nineteenth century physics. As for the physics of tomorrow, which Rhine says is a pointless speculation, we should remember, as Walker (**8**) points out, that our most valuable instruments of scientific progress are the working hypotheses which look forward by extrapolation from present knowledge in all promising directions. Without these prescient excursions into the unknown, growth would be arrested and science would die of senile sclerosis.

We are surprised also to read that the claim that human mentation as a vital function is "a metaphysical assumption" or "mere conjecture," and that psychologists have the vague notion that mind and body "are somehow fundamentally unified on some complex but wholly unknown physical basis," and that "this half-formulated materialism is taken for granted; it has not been subjected to experiment" (**5**, p. 197). The fact is that it would be possible to cite literally thousands of experiments devoted explicitly to this problem—experiments made by the most competent neurologists, physiologists, and psychologists we have. These experiments are of many kinds, using a wide variety of methods.

The most instructive of these experiments are based on the fact that all nervous action is accompanied by changes in electrical potential that can be localized and accurately measured. Using the same technique, a particular kind of mental work can be shown to be accompanied by similar changes in potential which are localized in specific regions of the body. The "brain waves"

recorded by the oscillograph of thousands of conscious persons have been studied under a great variety of conditions, and the effects of various kinds of mental experience are clearly seen in this permanent record of the related changes in electrical potential. This is only one among many other methods which have been used to prove that we think with our bodies, and with different parts of the body for different kinds of thinking.

THE SPIRITUAL LIFE

It is possible to mention here only a few samples of the evidence which justifies the conviction that man's spiritual life, in the ordinary vernacular meaning of that term, is a real and significant component of his natural life. For the biologist this means that all mental acts of whatever kind are as truly vital processes as are nervous conduction and muscular movement. This gives us a unitary system of scientifically acceptable principles for everything within the range of human experience.

A man has native capacities for spiritual culture which set him at the highest level of integrative and creative efficiency yet reached in cosmic evolution so far as we know. This physicalistic conception of human nature, which has been accused of degrading the dignity of man and destroying his spiritual values, really points the way to the most efficient measures for enhancing these most precious human treasures. By setting the spiritual life in vital operational relationship with all other domains of science and of human endeavor, we get the benefit of all the resources of the sciences to reinforce and guide into appropriate channels the efforts of the other agencies of spiritual culture—education, philosophy, art, religion, and all the rest. This is a magnificent achievement.

REFERENCES

1. GIBBS, F. A. The most important thing. *Amer. J. publ. Hlth*, 1951, **41**, 1503–1508.
2. HERRICK, C. J. Introspection as a biological method. *J. Phil.*, 1915, **12**, 543–551.
3. HERRICK, C. J. A biological survey of integrative levels. In R. W. Sellars *et al.* (Eds.), *Philosophy for the future*. New York: Macmillan, 1949. Pp. 222–242.
4. RHINE, J. B. *The reach of the mind*. New York: William Sloane, 1947.
5. RHINE, J. B. *New world of the mind*. New York: William Sloane, 1953.
6. SELLARS, R. W. An analytic approach to the mind-body problem. *Phil. Rev.*, 1938, **47**, 461–487.
7. THOMPSON, D. W. *On growth and form*. New York: Macmillan, 1942.
8. WALKER, R. Parapsychology and dualism. *Sci. Mon.*, 1954, **44**, 1–9.

(Received October 14, 1954)

Psychological Review
Vol. 62, No. 5, 1955

THE STIMULUS CONDITIONS WHICH FOLLOW LEARNED RESPONSES

CHARLES C. PERKINS, JR.

Kent State University

The further development of a reinforcement theory of learning, or, for that matter, of any other general learning theory, would seem to require a systematic analysis of the different procedures used in the experimental investigation of basic learning processes. The present paper will be limited to a part of this task, the analysis of the stimulus conditions which immediately follow occurrences of the response acquired or extinguished.

The approach taken here is to assume that any stimulus-response connection which becomes stronger is followed by a more reinforcing stimulus situation than alternate responses which tend to drop out. The frame of reference of this paper resembles that of Meehl (11) in his evaluation of the law of effect.

An analysis of the stimulus situations which follow the alternative responses in a particular context is assumed to disclose the relative reinforcing properties of these stimulus situations. Once the relative reinforcing properties have been determined by the observation of learning, they can be used to predict learning of other responses. This approach, one typical of explanations of trial-and-error learning, will be applied to other "types" of learning as well. I will then show that the results of the analysis seem to require a somewhat different treatment of conditioning from that represented by typical reinforcement theory (7, 12), contiguity theory (2, 5), or two-factor theory (13, 16). The analysis to be made below seems to clarify the distinction between trial-and-error learning and conditioning at the same time as it implies that common principles underlie all learning.

In the final sections I will (*a*) show how the analysis suggests an explanation of the development of observing responses under conditions which would not lead one to predict their acquisition from other viewpoints; and (*b*) suggest several direct experimental tests of the interpretation of learning outlined in this paper.

Since an analysis is more meaningful if terms are explicitly defined, we now turn to brief definitions of some of the terms and symbols used in this paper.

Definitions and Symbols

As the term is used in this paper, a *stimulus* is any situational variable which, if appropriate training were given, could come to be followed by a response which would not occur in the absence of this stimulus. Any stimulus which is an environmental condition present at the moment being described is an *external stimulus*.

An *internal stimulus* is the stimulus aftereffect of (*a*) an environmental condition which was present earlier (whether or not it remains at the moment in question), (*b*) a preceding response, or (*c*) a deprivation schedule.

A *situation* is the stimulus complex at a given moment in so far as it is specified by a report of the experimental procedure. Since no description includes all observations or measures that could be made, a situation (like a stimulus) is a generic term referring to a class of situations. However, unless otherwise indicated, it will be assumed that reports of experimental procedures

are complete enough to indicate (at least by implication) any systematic difference between situations which precede a correlated difference in subsequent events.

It should be noted that the above definition of a stimulus includes stimuli defined as changes in the environment. However, unlike the latter definition, ours urges an inventory of all situational variables. Such a listing of situational variables appears to facilitate the determination of response-evoking conditions as well as conditions affecting reinforcement.

The term *response* will be used to include the nonoccurrence of an activity as well as the occurrence of an activity. This usage seems preferable because "resting," or not making a given response, appears to be learned in the same way that an active response is acquired.

Our use of the term *reinforcement* is best explained in conjunction with a description of trial-and-error learning, to which we now turn.

TRIAL-AND-ERROR LEARNING

The basic working hypothesis of this paper is the assumption that whenever two responses occur to the same stimulus situation, and the tendency for one (R_A) to occur gains strength at the expense of the tendency for the other (R_B) to occur, the stimulus situation which follows the first (X_A) is more reinforcing than the situation which follows the second (X_B).[1]

Actually, to avoid circularity, X_A and X_B must refer to classes of situations. Perhaps the classes of situation most frequently contrasted in learning experiments are a context which includes hunger (X_c) and food (X_a), on the one hand, and the same context (X_c) and no food (X_b), on the other.[2] Hunger is, of course, just a shorthand way of labeling the aftereffect of a period of food deprivation.

Strictly speaking, the stimulus aftereffects of R_A and R_B cannot be identical, and therefore situations X_A and X_B cannot be identical either. However, the contexts following R_A and R_B may be equivalent; i.e., stimuli may have the same relative reinforcing properties in either context. For example, either right or left turns in a T maze are ordinarily learned about equally rapidly when reinforced by food or any one of a number of other reinforcing stimuli. Thus the context of goal box following a left turn is equivalent to goal box following a right turn.

When two situations are equivalent except that the first includes X_a and the second X_b, and R_A increases relative to R_B, then X_a (a stimulus, not a situation) may be said to be more reinforcing than X_b. However, it should be recognized that this does not necessarily imply that X_a would be more reinforcing than X_b in another context. Studies of drive discrimination (6, 9) provide perhaps the best known experimental results which demonstrate clearly that the more reinforcing stimulus (water as more reinforcing than food) in one situation (which includes water deprivation) is less reinforcing in another context (food deprivation). However, if we recognize the implied specificity of the context it is meaningful to speak of reinforcing stimuli as well as reinforcing situations.

[1] A further advantage of our definitions of response and of reinforcement is that the question of whether punishment weakens preceding response tendencies, or only leads to the strengthening of competing response tendencies, becomes meaningless. If R_A and R_B are two mutually exclusive response categories which include all possible responses, there is no way of discriminating between the strengthening of R_A and the weakening of R_B.

[2] Capitalized subscripts refer to situations, while stimulus components are designated by uncapitalized letters. Similarly, S_A or S_a refer to the situation or stimulus which precedes R_A, while X_A and X_a designate events which follow R_A.

When the expression "reinforcing stimulus" is used, it ordinarily implies that the situation plus this stimulus is more reinforcing than the situation without this stimulus. The implied description of situation minus stimulus is usually unambiguous. Similarly, a noxious stimulus ordinarily implies that the situation plus this stimulus is less reinforcing than the situation without this stimulus.

The above example of learning in a single choice-point maze is ordinarily considered a typical case of trial-and-error learning, and our analysis does not differ markedly from the usual explanation. However, after an examination of several special, atypical cases of "trial-and-error" learning we shall show that the same sort of analysis is quite adequate for an explanation of Pavlovian conditioning.

One technique for establishing a tendency for a striped-muscle response to occur in a given situation where it originally had little or no tendency to occur is the differentiation of a response (3). This technique consists first of reinforcing any responses which occur and have some similarity to the desired response; then, after the tendency to make similar responses is sufficiently strong, reinforcement is gradually restricted to those responses most similar to the response the experimenter wishes to establish. Later, after the desired response has occurred, reinforcement of all other responses is withheld.

This technique has probably been most fully described in connection with training pigeons to peck a key. However, it is a common procedure as part of preliminary training in other situations, and would appear to be applicable to the attaching of almost any striped-muscle response which the subject is capable of making to practically any stimulus situation.

In differentiation of a response, as in typical trial-and-error learning, the reinforcing stimulus is either presented or withheld; in other words, *all-or-none reinforcement* is used. However, *varied reinforcement*—Hull (8) has used the more descriptive but more cumbersome expression "correlated reinforcement intensities"—is another procedure which can also lead to the acquisition of a response which would not otherwise occur and thus could not be reinforced. In target shooting and many other motor skills the degree of reinforcement depends on the accuracy of the response. Therefore, the "better" or more accurate responses are always reinforced relative to the "worse" or less accurate ones.[3]

Use of varied reinforcement to establish new striped-muscle responses in animal subjects probably would not be a practical experimental procedure. The difficulty of providing a sufficiently distinctive dimension of reinforcing stimuli could not be easily overcome. This is perhaps the reason Hull applied correlated reinforcement intensities only to examples in which human subjects acquired responses which were strengthened by secondary reinforcement.

Three types of trial-and-error learning have been distinguished above. They are: (*a*) the most typical case, all-or-none reinforcement with the same class of responses reinforced throughout; (*b*) all-or-none reinforcement with gradual restriction of the responses reinforced; and (*c*) varied reinforcement, where the degree of reinforcement varies depending on the exact nature of the response. Since "new" responses can be acquired by either of the latter two procedures, the claim of some writers (10) that

[3] Actually, if level of aspiration is emphasized, learning of many skills can be adequately treated in terms of all-or-none reinforcement, in which case the learning would be a special instance of differentiation of the response. Probably both changes in level of aspiration and varied reinforcement are ordinarily involved in the acquisition of such skills by human subjects.

"new" responses are never acquired by trial-and-error learning does not fit the facts.

PAVLOVIAN CONDITIONING

It is the contention of the writer that the major difference between trial-and-error learning and Pavlovian conditioning is the nature of the reinforcing stimulus.

Since our basic premise is that responses followed by a more reinforcing stimulus situation are strengthened at the expense of those followed by a less reinforcing situation, we must look for the difference between the situation which follows marked salivation and the situation which follows little or no salivation in a Pavlovian conditioning experiment. Apparently there are no external stimuli which follow the salivary response to the CS but do not follow presentation of the CS and absence of the salivary response. Thus, the learning must depend upon reinforcement by an internal stimulus.

The only stimulus which systematically differentiates between the conditions immediately after presentation of the CS on trials on which a conditioned salivary response occurs and trials on which the CR does not occur is the stimulus aftereffect of salivation, or saliva-in-the-mouth.[4] Therefore, the occurrence of salivary conditioning when the US is food implies that a situation consisting of context, food, and saliva is more reinforcing than a situation consisting of context, food, and absence of saliva. The reinforcing stimulus, the stimulus component which distinguishes between the two situations, is saliva—an internal stimulus, not a part of the external environment.

Extinction of the salivary response when the US no longer follows the CS implies that a situation made up of context, absence of food, and absence of saliva is more reinforcing than a situation consisting of context, absence of food, and presence of saliva. Since salivary conditioning appears to be a general phenomenon, food plus saliva must be more reinforcing than food without saliva in a wide variety of contexts. Similarly, it appears that in many contexts the compound stimulus absence of saliva and absence of food is more reinforcing than presence of saliva and absence of food. The generality of these phenomena permits prediction of learning in a new context. Thus, our treatment is not circular.

A possible objection to the above treatment of Pavlovian conditioning is that the conditioned response has no initial tendency to occur and therefore cannot be reinforced. However, it can be demonstrated that such an objection does not apply if the following points are recognized: (*a*) some salivation occurs normally even if no food or "signals" of forthcoming food are present; (*b*) presumably there are fluctuations in rate of salivation from one trial to the next; and (*c*) varied reinforcement is involved in salivary conditioning.

The relation between rate of salivation and the reinforcing properties of the situation which follows is presumably similar to that shown by the solid line in Fig. 1. On the early conditioning trials the salivary response will be less than x (Fig. 1), but will vary between 0 and x (from points *a* and *b* above). During these trials the greater the amount of salivation to the CS the more reinforcing the situation which follows. A positive relationship between salivation and degree of subsequent re-

[4] For convenience, saliva or saliva-in-the-mouth will be used interchangeably with the stimulus aftereffect of salivation, even though there are probably also other stimulus aftereffects of salivation. It should be noted, however, that the physiological nature of stimulus aftereffects of an event (e.g., saliva in the mouth) need not be specified for postulation of stimulus aftereffects. The expression "saliva in the mouth" is used here on the assumption that it will facilitate the exposition for physiologically minded readers.

inforcement exists until rate of salivation surpasses *Opt.*, the rate of salivation which is followed by maximum reinforcement. Thus, rate of salivation will gradually increase until the mean rate is at *Opt.*, and since further increase in salivation is followed by a less reinforcing situation, rate of salivation will stabilize at point *Opt.*

The dotted line in Fig. 1 represents the reinforcing properties of stimulus situations which follow various rates of salivation when no food is presented, as on extinction trials. The process of extinction may be explained in the same manner as was conditioning.

This treatment of learning implies that a conditioned response is a preparatory response. This interpretation is strongly supported by recent work (1, 4, 14, 18) which rather clearly indicates that effective conditioning occurs only when the CS precedes the US. Instances of apparent backward or strictly simultaneous conditioning can all be explained as special cases of pseudo conditioning. Pseudo conditioning in turn appears to be only a special instance of the operation of stimulus generalization (19).

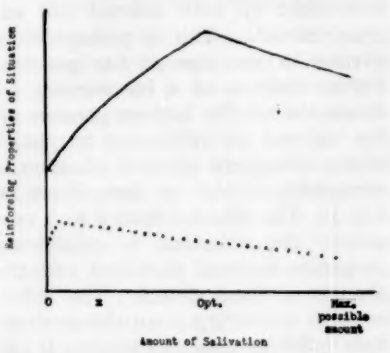

FIG. 1. The hypothetical relation between amount of salivation and the relative reinforcing properties of the stimulus situation which follows the CS on food-reinforced trials (solid line) and on extinction trials (dotted line).

TYPES OF RESPONSE MEASURES

The type of response measured is closely related to the type of reinforcement. As our treatment of salivary conditioning illustrates, smooth-muscle responses would seem to be acquired only through varied reinforcement by internal stimuli.[5]

Striped-muscle responses (and the responses of not engaging in certain striped-muscle activities) are typically reinforced by external stimuli as external stimuli are defined in this paper. However, there are some exceptions—as, for example, the response of not biting the tongue which is reinforced relative to biting the tongue.

A particularly important class of internal stimuli which determine the extent to which striped-muscle responses are reinforced is the stimulus aftereffects of activity or "fatigue" effects. Specifically, this class of stimuli provides greater reinforcement for inactivity than for activity.

If the response acquired is bar depression or a certain turn in a single choice-point maze containing food on but one side, and if only the occurrence or nonoccurrence of the response is recorded, internal stimuli (such as fatigue effects) need not ordinarily be considered. In such studies all occurrences of, say, bar pressing are followed by a more reinforcing state of affairs than any occurrences of not pressing the bar.

If pressure on the bar or running speed is the response measure, the relative reinforcing property of the situation following a particular response is

[5] Smith's ingenious suggestion (17) that smooth-muscle conditioning is an artifact resulting from trial-and-error learning of striped-muscle responses and unlearned bonds between stimuli from striped muscles and smooth-muscle responses is essentially a physiological theory which has no necessary relevance to the subject of the present paper. However, Smith's interpretation fits in very harmoniously with the behavioral principles advocated here.

determined by both internal and external stimuli. This is perhaps most obvious in the case of bar pressing. During learning of a bar-pressing response the relation between pressure on the bar and the reinforcing properties of the subsequent stimulus situation is presumably similar to that shown in Fig. 2. The distance from *a* to *b* represents the difference in reinforcing properties between click-food and the absence of these stimuli. The differences in reinforcing properties of situations following responses between 0 and *x*, and between *x* and *m*, result from kinesthetic and proprioceptive stimuli.

When running speed is the response studied, a similar analysis can be made. However, a detailed analysis of this type of learning is a good bit more complex and would require consideration of secondary reinforcement and/or anticipatory responses.

The definition of (relative) reinforcement used here has certain disadvantages which do not characterize an absolute (e.g., drive-reduction) definition. In particular, knowledge of the reinforcing properties of a stimulus in one context is not enough to predict the reinforcing properties of this stimulus in another context. However, even a drive-reduction theory implies as much; the drive (e.g., deprivation schedule) and goal object together are required for reinforcement. Thus, unless drive is specified in terms of physiological activity rather than antecedent environmental conditions, drive reduction implies that the reinforcing properties of a stimulus depend on contextual stimuli. Unless it has acquired secondary reinforcing properties, food is a reinforcing stimulus only if there is also a hunger stimulus.

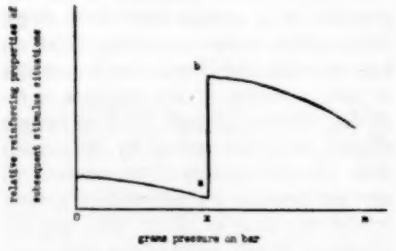

FIG. 2. The relation between pressure on a bar and the reinforcing properties of subsequent stimulus situations when *x* grams of pressure activate the relay and release a food pellet.

It thus appears that, for a molar behavior theory, stimulus patterning is an inescapable fact which affects not only the discriminative (perceptual) properties of situations but also their reinforcing properties.

THE ACQUISITION OF OBSERVING RESPONSES

Wyckoff (20) developed an ingenious technique which made possible the measurement of observing responses (R_o). His apparatus was a pigeon "Skinner box" in which there was a treadle as well as a key which was attached to the food-release mechanism. The apparatus was so wired that when the subject made the R_o of stepping on the treadle the key became either red or green.[6] When the subject was not standing on the treadle the key was white. Pecking at the red key was periodically reinforced by food; pecking the green key was never followed by food. Presentation of food was unrelated to the occurrence of R_o so that when R_o did not occur food would sometimes follow pecks to the white. The subjects showed a clear increase in R_o under these conditions, and a marked decrease when food presentation was no longer related to the color of the key. Wyckoff postulated that "information" has reinforcing properties and suggested, without elabora-

[6] Strictly speaking, R_o is not an observing response in this study or in Prokasy's. Following Wyckoff's usage, we use R_o to refer to any response which provides discriminative (positive and negative) stimuli which are not provided unless R_o occurs.

tion, that secondary reinforcement is probably involved. Thus, in effect, he described his findings but did not derive them from other principles.

Prokasy (15) ran rats in a single choice-point maze where either turn led to food on half the trials. All turns were followed by a 30-sec. delay before entrance to the end box. Either turn led to a black delay and goal box on half the trials, and a white one on the other half of the trials. On the consistent side one color (black) delay box was followed by access to food 100 per cent of the time, while the other (white) never led to food. On the inconsistent side either color delay box was followed by access to food on half the trials. Thus, R_o represents selection of the consistent side.

An increase in percentage of free turns to the consistent side was acquired by between 9 and 11 of the 12 subjects depending on the particular measure used. It should be noted that the two sides did not differ with respect to the amount of food received or the color of the goal and delay boxes. The acquired preference could only have resulted from the relation between color of delay box and food presentation.

Prokasy's findings follow logically from an analysis of the reinforcing stimuli following occurrence of a turn to either side.

Table 1 contrasts the stimuli which follow choice of the consistent side with those following choice of the inconsistent side, on food-reinforced and on non–food-reinforced trials, subsequent to the acquisition of anticipatory salivary responses.

Since each of the situations following responses to the consistent side is more reinforcing than the corresponding situation (same food condition) following a choice to the inconsistent side, the former choice will be acquired. A similar analysis can be used to explain Wyckoff's results.

DIRECT EXPERIMENTAL TESTS

The viewpoint outlined in this paper can most easily be subjected to direct experimental tests in connection with Pavlovian conditioning. While there are obvious technical difficulties, surgical techniques which would completely prevent salivary secretion from entering the mouth should lead in time to a decrease in the rate of conditioning and extinction, and perhaps to the eventual absence of any relation between salivation and stimulating conditions. Drive-reduction, contiguity, and two-factor theories would not seem to lead so clearly to such a prediction. However, this most certainly would not be a crucial experiment, since any general principle can be made consistent with any single experimental result.

If technical problems were overcome, several further experiments could be conducted involving the clamping of the salivary ducts so that the flow of saliva could be prevented whenever the experimenter wished. For example, the ducts could be blocked whenever the subject was placed in a highly distinctive locus and at no other time. If the subject had been frequently fed in this locus, we might expect that salivation would no longer accompany eating here,

TABLE 1

THE STIMULI SUBSEQUENT TO CHOICE OF EITHER SIDE IN PROKASY'S EXPERIMENT

Goal-Box Condition	Stimuli Following a Response to the Consistent Side (Correct "Anticipation" Possible)	Stimuli Following a Response to the Inconsistent Side (Correct "Anticipation" Impossible)
Food*	Food plus optimum (much) salivation	Food plus less than optimum salivation
No food*	No food plus optimum (very little) salivation	Food plus more than optimum salivation

*Food followed either response 50 per cent of the time.

as saliva would not enter the mouth, but instead salivary gland activity might inflate the gland to the point where it would be "painful." Presumably, the subject would eventually learn not to salivate while eating in the locus in which the ducts were blocked, but still continue to salivate while eating elsewhere.

Another possible experiment would involve a system for temporary storage of the subject's saliva in a container, and the injection of this saliva into the subject's mouth when a particular striped-muscle response is made. For example, it should then be possible to train a dog to stretch his right foreleg before and during eating in order to inject saliva into his mouth.

CONCLUSION

The emphasis on relative reinforcement suggested in this paper seems to have several distinct advantages, most notably that (*a*) it allows for a treatment of all learning by the same set of principles, without involving the complications which a drive-reduction theory necessitates;[7] and (*b*) it permits derivations concerning the acquisition of observing responses which are not instrumental in attaining a goal object from principles already demonstrated experimentally.

REFERENCES

1. BORASIO, G. An experimental investigation of backward, forward, and strictly simultaneous conditioning of fear. Unpublished master's thesis, Kent State Univer., 1952.
2. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
3. FERSTER, C. B. The use of the free operant in the analysis of behavior. *Psychol. Bull.*, 1953, **50**, 263–274.
4. FITZWATER, M. E., & REISMAN, M. N. Comparisons of forward, simultaneous, backward, and pseudo-conditioning. *J. exp. Psychol.*, 1952, **44**, 211–214.
5. GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.
6. HULL, C. L. Differential habituation to internal stimuli in the albino rat. *J. comp. Psychol.*, 1933, **16**, 255–273.
7. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
8. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
9. LEEPER, R. The role of motivation in learning: a study of the phenomenon of differential motivational control of the utilization of habits. *J. genet. Psychol.*, 1935, **46**, 3–40.
10. MAIER, N. R. F., & SCHNEIRLA, T. C. Mechanisms in conditioning. *Psychol. Rev.*, 1942, **49**, 117–134.
11. MEEHL, P. E. On the circularity of the law of effect. *Psychol. Bull.*, 1950, **47**, 52–75.
12. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
13. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
14. MOWRER, O. H., & AIKEN, E. G. Contiguity vs. drive-reduction in conditioned fear: temporal variations in conditioned and unconditioned stimulus. *Amer. J. Psychol.*, 1954, **67**, 26–38.
15. PROKASY, W. F., JR. The relative effects of two reinforcement procedures on a choice response. Unpublished master's thesis, Kent State Univer., 1954.
16. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century, 1938.
17. SMITH, K. Conditioning as an artifact. *Psychol. Rev.*, 1954, **61**, 217–225.
18. SPOONER, A., & KELLOGG, W. N. The backward conditioning curve. *Amer. J. Psychol.*, 1947, **60**, 321–334.
19. WICKENS, D. D., & WICKENS, C. Some factors related to pseudo-conditioning. *J. exp. Psychol.*, 1942, **31**, 518–526.
20. WYCKOFF, L. B., JR. The role of observing responses in discrimination learning. *Psychol. Rev.*, 1952, **59**, 431–442.

[7] This should not be taken as implying that the viewpoint expounded here is necessarily incompatible with a drive-reduction theory. In fact, if the viewpoint presented here is extended, it would seem to imply that all increases in response tendency result from "drive" reduction. However, the approach suggested in this paper seems more parsimonious and more easily related to experimental findings.

(Received November 15, 1954)

Psychological Review
Vol. 62, No. 5, 1955

CONSCIOUSNESS, BEHAVIOR, AND PERSONALITY

JOSEPH NUTTIN

University of Louvain

The concept of consciousness still remains associated with one of the most traumatic experiences in the life history of scientific psychology, i.e., introspectionism. The wound has not yet healed; consciousness still means for many people a menace to behavioral science. It would not be justified, therefore, to run the risk of arousing emotional reactions by reintroducing the subject if there were no hope that a revocation and a restatement of the problem in the present situation might entail some therapeutic effect and eventually help psychology in its development to greater maturity. Maturity, in fact, means getting in contact with the real behavioral world and facing the problems arising there.

Consciousness, now, is just what brings man in contact with the behavioral world. Contrary to what is generally believed, it should not be considered as necessarily related to introspectionism and the study of so-called "states of consciousness."

Perception and the Representation Theory of Consciousness

Some trends in present-day psychology emphasize again the role of cognitive and perceptual functions in behavior. This may be an important step in the direction of a sound concept of behavior. I see a danger, however, in the fact that, frequently, theories of cognition and perception are still working with concepts adopted from introspectionist psychology. The perceptual world is currently conceived as a "representation" *in* the subject of the world outside him. Thus, this representation is a kind of duplicate of the outside world: a duplicate existing *in* the subject. In a recent symposium on cognitive theory and personality functioning contributors repeatedly expressed themselves in such terms of representation. "Cognitive theory," Scheerer (4) argues, "assumes that the organism is in commerce with its environment through the medium of selective representation." Sweet (5) claims that "via the mediate of sensory processes, representations of the environment are obtained." This perceptual representation is considered a "perceptual response present in the organism."

In the field of cognitive and perceptual processes, as in behavioral science in general, it appears to be necessary indeed to return to what have been called the "naturalistic" or primary facts. If not, psychology may lose contact with the specific phenomenon to be explained, i.e., man behaving in his world of concrete life situations. From the psychological or behavioral point of view, nothing corresponds less to reality than to conceive of perception as a representation *in* man of something which exists outside him. Perception is a *direct* contact with the outside world *itself*. No medium, no duplicate, no representation, is interposed between the subject and his world.

This deep-rooted representational theory stems from philosophical and physiological problems. The psychology of sensation, for instance, studies the physiological processes involved in perception. From that standpoint, perceptual responses are in fact "organismic states which the organism carries about in the nervous system and

which *represent* the perceived object for me" (5). Philosophy, on the other hand, raises the problem of a *reality* existing beyond the object of cognitive processes. The object perceived is considered only a representation or a duplicate of that unknown reality as it is supposed to exist in itself, independently of the behaving subject. From the behavioral or psychological viewpoint, however, there is no doubt that man perceives the things themselves, that is, the world itself. Perception does not give him a *representation* of the world, but a direct *presence* of the world itself (1, p. 186).

The representational theory is responsible for introducing an erroneous concept of consciousness. Consciousness has been conceived as the enclosed world of inner representations in the subject. In fact, according to the representational theory, there is on the one hand the world of objective, observable things which transcends psychological activity; on the other hand, there is the psychological world of consciousness where the objective world is represented on an enclosed, inner stage. On this enclosed stage the inner representations of outside things are acting in a private and incommunicable play. Thus, representational theory is responsible for introducing a double stage in our world, and even a double world.

CONSCIOUSNESS: EXPOSURE TO THE OUTSIDE WORLD

As a matter of fact, the sciences of behavior are not directly interested in an incommunicable private world of consciousness. Psychology is right in turning to the observable world itself in which man really behaves. But turning to the objective, observable things does not mean denying consciousness. On the contrary, this world of objects, fellow men, and situations precisely *exists for us*, is *present* to us, by the very fact of man's "openness" or exposure to the outside things. This specific exposure constitutes consciousness.

Thus, the naturalistic fact to which behavioral sciences should return is that man, at any moment, *finds himself in a world* of things, fellow men, and meaningful situations. His behavior is nothing other than his commerce or dealing with this world of life situations. In other words, the primary fact to be taken into account in the conception and study of behavior is the fact that an outside world is *present to us*. This presence of an outside world consists in a *direct* contact with things and men outside ourselves. This means that man is not an enclosed structure in himself, but that he is essentially constituted as an "exposure" to an outside world. This concept of exposure and direct presence of the world to us is fundamental in this sense: it underlies the possibility of "meeting" people and things, and thus the possibility of *behavior* as a dealing with situations. Looking at something, speaking to an audience, attending a conference, going away or staying at home, feeling frustrated, etc. mean and imply being in the presence of an outside meaningful world.

This presence of the outside world to us is not to be conceived, of course, as a physical fact (physical neighborhood, for instance). A tree or a room may be situated in the same physical environment as we are, and nevertheless not deal with this environment as with their "world"; they do not manifest that a world is present to them or that they are living in a world. Their responses are limited to physical or chemical reactions to stimuli. Therefore, the presence of a meaningful world to us is a psychological or behavioral fact.

Consciousness in its primary sense

means precisely this exposure of the organism to the world, or in other words the direct presence of a world to us. The very fact, for instance, that an audience is present to me, or exists for me, as a part of my world or situation at the moment I address them, derives primarily from the fact that I am conscious. Consciousness means exposure to the outside world, not enclosedness in the "inner self." Moreover, consciousness or the fact that man finds himself in the presence of an outside world is not to be considered as a "private" affair of inner life. On the contrary, it is an object of direct perception. Seeing a man speaking with other people or driving a car, I perceive the fact that streets, houses, and people are present to this man, that they exist for him, that he is aware of them, that his behaving organism is open to them, i.e., that he has consciousness.

In order to avoid any misunderstanding, it should be stressed that recognition of this primary fact of consciousness as exposure to the outside world does not contradict the theory of perception as a selective and constructive process. The concrete content of the behavioral world present to us is a selective construction of the functioning psychophysiological personality. But no specific content of a perceptual world could ever be built up if there were not this fundamental property and activity of the organism which is called consciousness. In other words, man's exposure to the world is not the result or the total sum of a series of concrete perceptual contents. On the contrary, selective and constructive perception can only be conceived as a continuous activity of specification of the outside world to which man finds himself exposed. This fundamental fact of exposure as such is what is meant by consciousness. Without the fundamental capacity of the organism to be open or exposed to an outside world, perceptual processes would be reduced to internal organismic changes. *Lack* of consciousness, therefore, means enclosedness in the realm of physical and biochemical interactions. Consciousness is psychophysiological activity underlying all psychological or behavioral processes, insofar as they are intrinsically directed toward objects. Even the so-called unconscious motivations and behavior patterns presuppose consciousness in this fundamental sense.

Psychological and physiological sciences, of course, have to investigate the processes involved in the subject's exposure to the world. But the scientific examination of these processes should not destroy the fundamental psychological fact that a world is present to us, or exists for us, and that man's behavior is meaningful dealing with this world. The purpose of scientific psychology is to explain and to account for this fact, not to replace it, for greater convenience, by underlying but different processes. In other words, the underlying processes discovered by scientific investigation should ultimately be related to that unique behavioral phenomenon to be explained. The concept of consciousness does not give a solution to any scientific problem about behavior, but it is necessary for an adequate conception of behavior itself, and for a realistic approach to behavior problems.

Thus, the introduction of the notion of consciousness does not mean going back to introspectionism. On the contrary, it simply means the recognition that the essential characteristic of behavior itself is its being related to a meaningful outside world with which the subject finds himself presented. We also should not forget that scientific research itself is a form of behavior, i.e., a form of dealing with the world we are living in. Communicating sci-

entific results is another form of behavior. Both activities imply what we have called consciousness.

THE VICARIOUS PRESENCE OF THE WORLD IN CONSCIOUSNESS

Until now, we have considered consciousness as putting us in direct contact with the outside world and with the external behavioral facts. The problem of consciousness, however, becomes complicated by the fact that the world can be present to us in different ways. Man is able to find himself in the presence of an object or a situation which is not immediately presented to him from the outside. This is what I call the *vicarious presence* of an object. This possibility of vicarious presence of things to us, via images, concepts, and symbols, is a capital behavioral fact. It enlarges to an enormous extent the world in which man is living and, consequently, the possibilities of behavior. By this vicarious presence man's dealing with situations goes far beyond the limits of responding to actually present things. Man is, for instance, *thinking* about them; i.e., he is dealing with them in a cognitive way. Thinking and imagining are in fact nothing else than other ways of going along with the world and dealing with its situations. This inner, cognitive dealing with situations is intrinsically directed to, or projected toward, an outside world and it is, as a rule, preparatory or compensatory to effective external behavior. In some circumstances—as is the case, for instance, in science—thinking about things actually constitutes the effective form of behavior itself. In fact, man finds himself also in this very particular situation that the world is perceived as a problem to be solved, i.e., to be dealt with in a cognitive way.

MAN'S PRESENCE TO HIMSELF

Man not only finds himself in the presence of an outside world; he may also be presented with himself. His own activities may become for himself the objects of his perception and cognition, as is the case in introspection. In this way man finds himself presented with his own activities, thoughts, and feelings about the world: he is reflecting on himself, on his activities, and on the content of his spontaneous or direct consciousness. This inner presence of man to himself and the reflection upon his own activities may be called *reflective consciousness* or consciousness raised to the second power.

Reflective consciousness is another way in which man escapes "enclosedness" in himself. By the very fact of being able to reflect on himself, man never finds himself completely enclosed in the actual phase of interacting with the environment. Reflecting on what he is and what he does, man "goes out" of himself, transcends the spatial and temporal determinations of the self by judging and evaluating himself and his activity.

At the same time, man's presence to himself enables him to look at himself and at the world from different points of view and to reshape at any time, to a certain extent, his own personality. It enables him to intervene, intentionally, in the course of his own development in order to restructure his relationships with the environment. This possibility seems to be implied in some phases of the therapeutic process (3, p. 422).

Once more, this highest function of consciousness broadens man's world and the self in a considerable way. It frees him from enclosedness in his own subjectivity and enables him to look at things and at himself from an "objective" point of view. This possibility is the basis of such forms of behavior as

science and intentional efforts of self-improvement. Thus, consciousness in its different functions frees man from enclosedness in his own organism, in the immediate physical world, and in his own subjectivity. While some degree of exposure to the world seems to be implied in all behavior, the higher functions of consciousness only manifest themselves in specific forms of human behavior (language, culture).

CONSCIOUSNESS AND BEHAVIORAL SCIENCE

As a matter of fact, the vicarious presence of things and man's reflective consciousness of his own activity take place on the level of inner representation. These more complicated functions of consciousness, however, presuppose consciousness in its primary sense of direct presence of an outside world to man.

It is essential for the science of behavior not to separate representational consciousness from direct exposure to the world, as it is important not to separate the representional way of dealing with situations from effective behavior. The representional level of consciousness is not to be studied in itself, as opposed to real behavior, i.e., as a world of "existential" qualities of inner experience. Clinical psychology and projective techniques have shown how closely related these representations and images are with man's real behavior in the world he is living in. Therefore, the introspective data of reflective consciousness and the vicariously present world on the level of representation are to be studied in their functional relationship with behavior as a whole. Even for clinical psychology it is important to base its insights on a behavioral theory in which the representional and operational levels of behavior are functionally related.

On the other hand, the meaningful world itself as it is present to us by direct consciousness cannot be separated from behavior. In fact, the world of meaningful things and situations is a crystallization or condensation of behavioral patterns. A physical object or a part of the surrounding space becomes a desk, a phone, or an office insofar as a whole network of behavioral dealing, and possibilities of dealing, are involved in it (2, p. 447). In this way, consciousness in the objective sense of "the world present to us" is crystallized behavior.

Moreover, man's conscious presence in the world is a *behavioral* presence. This means that, for the human personality, finding itself in the presence of a world implies and entails, necessarily, behavioral dealing with the world. In this sense, consciousness or exposure to the world is already *implicit behavior*, while behavior itself is externalized consciousness or *realized* presence to the world. In any case, instead of opposing consciousness and behavior, one should realize that no human realities are more inseparably united.

CONSCIOUSNESS, PERSONALITY, AND MOTIVATION

Finally, I have to indicate briefly the relationship between consciousness and the concept of personality. Psychologists as a rule conceive of personality as an internal organization of traits, attitudes, aptitudes, and behavioral consistencies. In fact, personality is such an internal organization; but the most essential characteristic of personality is lacking in such a definition. Personality is essentially a structure which goes beyond its internal organization. Its most characteristic feature consists in the fact that an outlook on and an exposure to the world is included in its own constitution. Personality is a way of being and behaving *in* a world that *exists for* the self. The outlook

on a world or life situation is an essential constituent of personality itself. In this sense, the fundamental structure of personality is an "ego-world unity." Nothing could more profoundly mutilate the concept of personality than neglecting this aspect of exposure to the world which is an essential component of personality as such. This exposure, as I have said previously, is just what is meant here by consciousness. If, as Lewin says, it is true that the personality governing behavior is a part of the psychological field, it must not be forgotten that the psychological field itself, in the general sense of "living in a world," enters into the make-up of personality.

Personality conceived in this way as an ego-world structure provides us with a sound basis for a theory of human motivation. Motivation theory today tends, so to speak, to situate and localize needs in the organism itself. But from the behavioral point of view, needs are to be conceived as basic patterns of organism-environment or ego-world relationships which actually are required for organism or personality functioning. In fact, personality is made up of different types of such relationships, not only on the level of biochemical organism-environment interaction, but also on the level of cognitive and affective ego-world relationships. While the basic patterns of biochemical interaction—and thus the basic drives in this field—are preformed in the physiological constitution of the organism, the cognitively elaborated ego-world relationships are more personal in origin and nature. Motivations on this level of behavior present themselves as an infinite variety of ego-world relationships in states of tension. Basic needs on this ego level are to be conceived as certain general and significant trends which eventually can be discovered in, and abstracted from, this variety of ways in which the individual tries to insert himself in his world. Such a general trend might be called "need for self-realization," as this "significance" is discovered in, or abstracted from, the innumerable concrete forms of man's behavioral dealing with the world of his fellow men and life situations.

The patterns of ego-world relationships which constitute the infantile personality will be different from the adult ego-world structure. The same significant trends, however, may be discovered in both. In that way, there may be infantile and adult forms of the same basic need, as there are infantile and adult patterns of ego-world relationships. There is no reason why the infantile form should be considered the *basic* one, and the adult pattern the derived or secondary need. Both are rather to be conceived as varieties of a basic type of ego-world relationship which the individual needs in order to maintain and develop the functional organism-environment and ego-world structure which constitute personality at a given stage of its development.

SUMMARY AND CONCLUSIONS

Psychologists who work in the fields of cognition and perception still frequently consider the perceived world as an internal representation of a real external world. This representational theory is strongly rooted in psychology, but it originates in a philosophical speculation which raises the question of a reality underlying the objects of cognition. From the viewpoint of the psychological study of behavior, the perceived world is by no means a representation *inside* of man (or inside consciousness) of some object whose real existence lies outside of him. The perceived world is the real world itself, and perception consists of direct con-

tact with the environment. For this reason the notion of consciousness, as it has been developed within the framework of introspective psychology, is in need of revision. Consciousness does not refer primarily to an internal world made up of representations; it consists of a direct awareness and of an immediate *presence* of the world itself.

Consciousness therefore signifies, before all else, *exposure to the real behavioral world,* and not confinement in an internal world of representation. Consciousness establishes a man as a being who lives in the world, and thus consciousness is an essential aspect of behavior.

The world, however, can also be present to the behaving subject in a *vicarious* way, via representation. Moreover, man can be present to himself by reflective consciousness. These "higher" functions of consciousness also free man from specific forms of enclosedness.

As for the concept of personality, nothing is less in conformity with reality than to conceive of personality as a pure internal structure of traits, aptitudes, etc. Personality signifies a specific modality of existence and of activity which is characterized by the fact of involvement in a situation or in the world. The fundamental personality structure is therefore an ego-world unity, and not an internal organization of an ego. The psychological field itself, more basically, enters into the make-up of personality. Motivation also must be looked at in relation to the fact that man lives in a meaningful world. In fact, basic needs are general types of organism-environment and ego-world relationships which actually are required by the functioning of organism and personality.

The introduction of the notion of consciousness solves no concrete scientific problem regarding behavior, but it is necessary to appeal to the concept in order to formulate and to come to grips with the problems of behavior and of personality in a realistic fashion.

REFERENCES

1. HUSSERL, E. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie.* Halle: Max Niemeyer, 1913.
2. NUTTIN, J. *Tâche, réussite et échec: théorie de la conduite humaine.* Louvain: Editions Universitaires de Louvain, 1953.
3. ROGERS, C. R. Significant aspects of client-centered therapy. *Amer. Psychologist,* 1946, 1, 415–422.
4. SCHEERER, M. Personality functioning and cognitive psychology. *J. Pers.,* 1953, 22, 1–17.
5. SWEET, A. L. Some problems in the application of cognitive theory to personality functioning. *J. Pers.,* 1953, 22, 41–52.

(Received November 18, 1954)

Psychological Review
Vol. 62, No. 5, 1955

THE DIFFERENCE BETWEEN *Q* AND *R*

JACK BLOCK

Institute of Personality Assessment and Research, University of California, Berkeley

While a good deal has been said by now about the merits and inadequacies of the *Q* approach to personality data, it seems to the present writer that a misunderstanding still persists as to the relationship between *Q* and *R* results. It is still widely believed that the alternative approaches of correlating *variables* and correlating *persons* produce identical, or rather fully convertible, results. A recent paper by Eysenck (4) iterates this point.

It is the purpose of this note to specify the conditions where the equivalence between *Q* and *R* will exist and the conditions where this equivalence will *not* occur. The controversy which has arisen devolves from differing initial assumptions about the nature of personality rather than being an issue in the realm of matrix algebra.

Let us suppose we have by some means finally established or decided upon the *n* basic variables in terms of which personality is to be described or measured. Shall we find at this point that for all possible subgroups of individuals these variables produce equivalent correlation matrices? Or shall we find that the variables interrelate differently in different individuals or "types" of individuals with fundamentally different correlation matrices resulting?

At our present stage of knowledge, our research is guided and our results to some extent limited by the assumption we choose to make as to this ultimate finding. But even now, evidence is available bearing upon this question. Many findings exist wherein the interrelationship of variables differs, in a reliable fashion, as a function of the particular sample investigated. For example, Eysenck (5, p. 205) reports that for Communists, the correlation between direct and indirect aggression is − .94 while for Fascists, the correlation between the same two variables is + .61. For a Neutral group, the correlation is − .64.

If the assumption that variables interrelate differently within different types of individuals is held, therefore, as most consonant with the psychological reality, the way is opened for the possibility of differences between the factors obtained from *R* and *Q* analysis. For *the results of* Q *technique are not convertible into the results of* R *where the subject population providing the data for analysis is heterogeneous with respect to the modes of interaction between variables. Where homogeneity of interaction between variables exists for all individuals within the subject sample, results from* Q *and* R *are transposable.* By heterogeneity of interaction between variables is meant the situation where the pattern of correlations between variables for one subgroup within the total sample differs reliably from the patterns of correlation derived from other subgroups. That is, the variables, X, Y, Z, and so on interrelate differently in one subgroup of subjects than they do in a second subgroup.

Within a sample of subjects homogeneous with respect to such external criteria as age, sex, intelligence, socioeconomic status, and so on, there may still exist a heterogeneity of interaction between variables. *This is particularly true where the variables being interrelated are derived from the personality domain.* Different "styles" of personality organization are known to exist

even within samples homogeneous with respect to conventional criteria.

Some illustrations will help make this point clear. The variable "impulsivity" correlated .54 with "suggestibility" in one population (Air Force captains), and − .50 with "suggestibility" in another population (male graduate students). These differences in correlation are highly reliable. If a sample of subjects had been drawn from each population and these samples combined, the correlation between "impulsivity" and "suggestibility" when computed for the combined sample would be zero, or tilted in one direction or the other, as a fortuitous function of the number of individuals included in each subgroup. An *R* factor analysis of each of these subgroups separately or of the combined sample would yield three widely divergent results which could not be reconciled and which in any event would not be revealing of the various psychologically possible relationships present. On the other hand, if persons are correlated in the above illustration (assuming that such correlation is metrically feasible), four types of individuals emerge: those individuals both impulsive and suggestible; those individuals who are neither impulsive nor suggestible; those individuals who, while impulsive, are not suggestible, and finally, those individuals who are not impulsive but are suggestible. Depending upon the continuity or discontinuity of the relationships involved, individuals intermediate to these polar types may be found. These varying personality patterns would seem to have sufficient behavioral implications to warrant their discernment. As indicated above, some or all of these may well be lost by *R* analysis.

Generalizing from the above illustration, it may be shown that where a matrix based upon correlated variables contains zero-order correlations, thus providing little or no possibility of deriving factors, the same data arranged to provide correlations between persons can result in perfect separation of individuals into type factors.[1] While such extreme situations do not exist in nature, approximations to it to a greater or lesser degree do exist—to an extent we have tended to neglect in our investigations in the past.

Another illustration, this time less hypothetical, of the differences in results obtained by *R* and *Q* analyses will further emphasize our basic point. One hundred military officers were assessed over a three-day period and the evaluations of staff psychologists of each subject were codified by means of the *Q*-sort method of description.[2, 3] The staff *Q* descriptions of each officer were correlated person by person and then cluster analyzed by Tryon's method (8).[4] The first four clusters which emerged seemed, in terms of their defining characteristics, to be arranged along an impulse-control dimension. At one end of the underlying dimension were the dashing individuals who cannot or do not modulate their impulses and needs. At the other end of the continuum were the con-

[1] Several years ago, Meehl (6) made essentially this same point in an article on the discrimination possibilities afforded by configurally scoring personality items.

[2] This research was supported in part by the United States Air Force under Contract AF 18 (600)–8, monitored by Technical Director, Detachment No. 7 (Officer Education Research Laboratory), Air Force Personnel and Training Research Center, Maxwell Air Force Base, Alabama. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[3] Personal views or opinions expressed or implied in this publication are not to be construed as necessarily carrying the official sanction of the Department of the Air Force or of the Air Research and Development Command.

[4] The differences between factor and cluster analysis do not affect the argument being advanced here.

stricted, defensively repressive individuals. Occupying a middle stretch along the dimension were the comfortably conscientious individuals, relaxed and appropriate in their behavior. A group of rigidly conforming men fell between this latter group and the utterly constricted group.

Now while the officers defining the two poles of the dimension differed strikingly in their ratings of *rigidity, fluency, breadth of interest,* and so on, they were alike in being described as *sarcastic, guileful,* and *lacking in dependability.* It is thus obvious that a correlation between, for example, the variables "sarcasm" and "rigidity" computed over the total sample of 100 will be affected by the differing ways in which the variables are related for the several groups being improperly combined. In point of fact, this correlation is $-.11$, clearly insignificant and uninformative. As other data show, the meaning of sarcasm in the diametrical clusters is different. For the impulsive individuals, sarcasm is seen as an additional expression of their unmodulated spontaneity. For the threatened, constricted persons, sarcasm functioned as a means of externalizing inadequacy feelings and projecting hostility. The nature of these relationships would have been lost if the procedure of correlating persons had not been used.

The argument we have here been advancing is by no means new. Stephenson (7), of course, has propagated it for years. Cronbach (3) also has called attention to the possible difference between *R* and *Q* results. Even Cattell (2), who has vigorously argued against *Q*, says that "*Q technique* has its chief use as a classificatory device for finding the subpopulations in a nonhomogeneous population . . ." (2, p. 502).[5]

In calling attention to the difference between the *R* and *Q* approaches, the present writer does not intend to encourage pell-mell abandonment of *R* and total acceptance of *Q*. There are a number of difficulties involved in formulating data suitable for *Q*, while *R* still remains the method of choice for data situations where the sample being evaluated is homogeneous with respect to "divergent" variables. The purpose of this note will fully have been realized if merely it is recognized that *R* and *Q* are not always fully convertible. It is also suggested, along with Beck (1), that at the present stage of personality research, correlating persons deserves a greater emphasis than it has received in the past.

REFERENCES

1. BECK, S. J. The science of personality: nomothetic or idiographic? *Psychol. Rev.,* 1953, **60**, 353–359.
2. CATTELL, R. B. The three basic factor-analytic research designs—their interrelations and derivatives. *Psychol. Bull.,* 1952, **49**, 499–520.
3. CRONBACH, L. J. Correlations between persons as a research tool. In O. H. Mowrer, *Psychotherapy: theory and research.* New York: Ronald, 1953. Pp. 376–388.
4. EYSENCK, H. J. The science of personality: nomothetic! *Psychol. Rev.,* 1954, **61**, 339–342.
5. EYSENCK, H. J. *The psychology of politics.* London: Routledge and Kegan Paul, 1954.
6. MEEHL, P. E. Configural scoring. *J. consult. Psychol.,* 1950, **14**, 165–171.
7. STEPHENSON, W. *The study of behavior.* Chicago: Univer. of Chicago Press, 1953.
8. TRYON, R. C. *Cluster analysis.* Ann Arbor: Edwards, 1939.

(Received December 6, 1954)

[5] Incidentally, it is interesting to note that although Eysenck (4) cites this article as supporting the equivalence of *R* and *Q*, Cattell does not definitely commit himself on the question of transposability. Thus, he reports that Burt "*claims* to have demonstrated [the transposability] . . . and *some* statisticians agree with this demonstration" (2, p. 503, italics mine).

Psychological Review
Vol. 62, No. 5, 1955

SOME PROBLEMS IN THE THEORY OF VIGILANCE [1]

JAMES DEESE

The Johns Hopkins University

In recent years there has been a revival of interest in problems that center around the classical notions of attention and set. For the most part, this revival is the result of practical concern with the use of human observers as monitors under conditions of monotony. This rather applied approach, however, has not meant that theoretical problems have been neglected. The British school (2, 3, 10) has been concerned with the theory of monitoring (or, as the British seem to prefer, watch-keeping) almost from the beginning of experimental work on the problem. Theory that has arisen from the current work has tended to ignore the classical problem of the description of subjective states in attention, and has centered around concepts useful in describing the relationship between efficiency in detecting signals and various environmental and organismic conditions.

Largely due to the influence of Mackworth (10), the words "attention" and "set" have been replaced by the term "vigilance." These words—attention, vigilance, and set—by no means refer to the same things, but they have common origins in a class of psychological problems, and it is a fruitless and arbitrary task to attempt to distinguish clearly between them. Mackworth has borrowed the term "vigilance" from Head, and he uses it to mean a high state of readiness to perform adaptive and purposive acts. In operational practice, the term is anchored to such derived measures of performance as probability of response to a given stimulus or inferred thresholds for some stimulus. Vigilance is an excitatory state, and in opposition to it, Mackworth proposes an inhibitory process analogous to the inhibitory processes discussed by Pavlov. Mackworth uses his inhibitory construct in a theoretical description of a general decline in ability to detect signals, a decline that he finds for relatively short periods of visual search. These declines occur when the conditions of search can best be described as monotonous. The major justification for the identification of such a decline in ability to detect with the growth of an active inhibitory state seems to be in the fact that Mackworth has observed an analogy to the phenomenon of disinhibition.

The present paper is concerned with further specifications of a concept of vigilance. The most basic assumption involved in the present theory is that the maintenance of a given level of vigilance in an observer depends to some extent upon stimulus events extrinsic to the observer. Much of this paper will consist of a discussion of alternative hypotheses about the exact nature of the stimulus events sufficient for the control of vigilance. The main emphasis in the present account will be upon excitatory properties, and with respect to the excitatory properties of vigilance there will be little in conflict with Mackworth, perhaps because Mackworth leaves vigilance largely unspecified. The present account will ignore inhibitory properties, except for a brief section intended to show some of the

[1] The research presented in this paper was performed under Contract No. AF 33(038)-22642 between the Institute for Cooperative Research, The Johns Hopkins University, and the Psychology Branch of the Aero Medical Laboratory, Directorate of Research, Wright Air Development Center.

advantages that result from minimizing the importance of an inhibitory construct.

Before presenting the hypotheses associated with the present conception of vigilance, however, it is necessary to mention briefly the operational framework within which the concepts are to be evaluated. The experimental problem to which the present notions are addressed is one in which observers search for small signals, at near threshold value, the location and exact time of occurrence of which cannot be predicted by the observer. The data (6) reported in the present account come largely from studies of simulated radar watching on a PPI-type cathode-ray tube, though except for special conditions noted below, it is hoped that these studies provide fairly general prototypes for many conditions of search. Certainly the degree of agreement between Mackworth's (10) studies of the clock test and the simulated radar test suggests a fair amount of generality to the experimental situations discussed in the present study.

TWO ALTERNATIVE HYPOTHESES

The first experimental problem to which the present notions were applied was that provided by varying the temporal distribution of signals.[2] It quickly became apparent that there were at least two hypotheses about the mode of action of the temporal distribution of signals upon detectability. Experimental evidence has tended to favor one of these over the other, but since the matter is by no means closed, a brief discussion of each hypothesis is given in the following paragraphs.

The first hypothesis was named the "reinforcement" hypothesis, for the good reason that its fundamental mode of action was built on an analogy to the role of reinforcement in learning. Let us see how the analogy applies to the general problem of search outlined above—that is, to search for small, near-threshold signals which occur at random times and places.

At the outset of a period of search or watch-keeping we may assume that the observer operates at some specified level of efficiency, and, for the moment, we may assume that this level is constant from one period of search to another. We can measure this efficiency in various ways, but for the present let us consider the measure provided by calculating the percentage of signals detected in some period of time. This can be considered as a rough approximation to the momentary probability of detection.

The observer's initial level of detection is determined by a host of factors such as motivation, skill, and the more specific receptor factors. We shall not be concerned with these things except to assume that they vary in a nearly random way from observer to observer and from occasion to occasion. After a probable warm-up period (6), it is assumed that the observer's ability to detect (probability of detection) will go through a simple course of linear decline. Rather than introduce an unnecessary inhibitory construct, we will assume that this represents a dissipation in the initial excitatory state of vigilance. Later it may be fruitful to relate this excitatory state to other determining states, such as motivation, but for the present problem we may leave these relationships unspecified.

It is apparent that we do not always get simple linear declines in probability of detection. This brings us to the critical assumption in the reinforcement hypothesis. The theory assumes that the occasional signals which occur in the

[2] The term "signal" is used throughout this paper to refer to those stimuli that the observer is instructed to detect.

observer's field of search determine the future course of his detectability. Specifically, it is assumed that either the *occurrence* or the *detection* of a signal leads to an immediate return of the observer's state of vigilance to its initial level or to a level determined by a change in some other experimental variable.

A precise statement of this assumption depends upon whether one indicates that the mere presence of a stimulus or its actual detection (as indicated by a verbal response) changes the level of vigilance. There are reasonable arguments that could be advanced for the view that the signal need not be detected in order to be reinforcing; it has long been suspected that observers maintain too high a criterion for admitting detection when it occurs, and there seems to be some ground for assuming "unconscious," i.e., unrecognized, detection of signals. Fortunately, it turns out that for certain purposes it does not make too much difference which version of the assumption we choose; so for present purposes we may phrase the hypothesis in terms of occurrence rather than detection. In an exact statement of the hypothesis, this assumption leads to more elementary mathematics.

One simple prediction which stems from the reinforcement hypothesis is that signals close together in time will, on the average, have a higher probability of being detected than signals spread over longer time intervals. This hypothesis has been tested in a simple experiment (6), in which observers reported the appearance of small, dim signals presented on a cathode-ray tube in a simulated radar-viewing test. The targets were randomized with respect to locus on the tube. With an important exception, they were also randomized with respect to time. The temporal distribution had the restriction that the average rate of signal appearance was specified and certain maximum and minimum times between signals were imposed. Four different rates were used, a rate of 10 signals per hour, one of 20 per hour, one of 30 per hour, and one of 40 per hour. The percentage of targets detected under each of these rates was measured in half-hour intervals over periods of search three hours in duration.

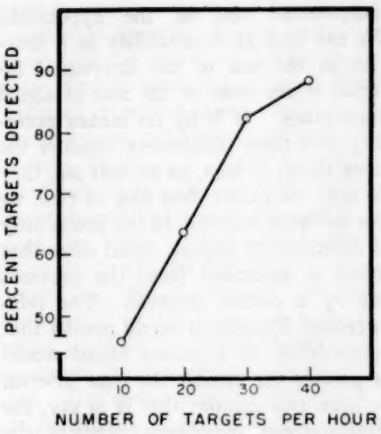

FIG. 1. Per cent signals (targets) detected as a function of the average rate at which targets appeared on the display. The points are averages for four subjects each on four three-hour periods of search.

The results of this experiment can be seen in Fig. 1. It is apparent that the hypothesis seems to be in accord with the facts. Detectability is higher when signals appear more frequently. Several subsequent experiments (6) verified this basic finding; signals appearing at a higher rate seem to be more readily detected. Furthermore, these experiments are in substantial agreement with those performed by Jenkins.[3]

However, another way of looking at the same data provides an additional,

[3] As reported in an unpublished paper by H. M. Jenkins, read in a Symposium on Vigilance before the American Psychological Association in New York, 1954.

independent test of the hypothesis. We can look at detectability as a function of the size of the intersignal interval rather than of the rate of signal appearance. It is by no means necessary that these parameters measure the same thing; indeed, as we will see, they do not. So rather than look at rate, we can see what happens to the probability of detection for a given signal when that signal is separated from the previous one by a certain interval. The reinforcement hypothesis would predict that detectability of a second signal would be greater the smaller the time interval between two signals; that is to say, the reinforcement hypothesis predicts the same thing for both rate and intersignal interval.

An exact test of this hypothesis demands that we consider the problem of whether it is the occurrence or the detection of a signal that has the postulated action of reinforcement. If we assume that the mere occurrence of a stimulus has a reinforcing effect, we would count intersignal intervals as the time between successive signals presented to the observer. If, on the contrary, we assume that it takes a detected (reported) signal to have the reinforcing effect, we would count the intersignal interval as the time from the last signal detected to a given signal. We have tabulated the data both ways, and while the outcome is not completely unequivocal (6), it certainly is true that the hypothesis is not supported. That is to say, intersignal interval either has *no* effect upon probability of detection (within a given rate) or the effect is in a direction *opposite* to that predicted by the hypothesis. In one experiment there was the suggestion that probability of detection could increase with intersignal interval.

In any case, the experiments from the author's laboratory, as well as analysis of Mackworth's experiments (10), fail to support the reinforcement hypothesis on this point. In view of this fact, the reinforcement hypothesis was abandoned and a new hypothesis explored. Like the reinforcement hypothesis, this hypothesis stems from the notion that a basic excitatory state of vigilance is maintained by conditions of external stimulation. This hypothesis, however, involves a different set of assumptions about how the excitatory state is maintained. These assumptions are embodied in what we may call an expectancy hypothesis, to keep the analogy to learning theory. The hypothesis states that (*a*) the observer's expectancy or prediction about the search task is determined by the actual course of stimulus events during his previous experience with the task, and (*b*) the observer's level of expectancy determines his vigilance level and hence his probability of detection.

Neither of these is a particularly obscure or difficult assumption. Extrapolation from the present behavior of the environment to its future behavior is a common enough activity of man. The accuracy with which observers can discriminate the rate of occurrence of signals does not become a serious question, since the hypothesis is not stated in a way to predict the exact level of detection. The hypothesis simply states that the feedback from the search task determines what the observer expects from further participation in the task in a simple proportional relationship and that his vigilance will vary accordingly.

With this notion of expectation determining output, it is easy to see that one arrives at the same prediction with respect to detectability and average rate of signal appearance as that made by the reinforcement hypothesis. As long as the average frequency of targets remains low, the expectancy of detecting one is low, and when frequency is high, expectancy will be high. With respect

to the intersignal interval question, however, the expectancy hypothesis and the reinforcement hypothesis part company. In the form outlined above, the expectancy hypothesis merely predicts that the detection for a given signal should be determined by the average of all signal intervals before the one in question. Thus the probability of detection for any given signal in a series should be the same as that for any other signal, assuming that the distribution of signal intervals is not some nonrandom function of time.

FURTHER DEVELOPMENT OF THE EXPECTANCY HYPOTHESIS

Any further prediction that the expectancy hypothesis might make about the effects of intersignal interval would stem from an extension of the first assumption of the hypothesis. The first assumption of the expectancy hypothesis states that the observer's prediction about the future course of stimulus events in the field of search is determined by the actual course of stimulus events in his past experience. Thus the assumption demands a linear relationship between the average of some selected aspect of past events and vigilance with respect to events at any given moment. If the task is invariant with time, vigilance will change little after the average level of expectancy has been reached. It is probable, however, that such an assumption demands that the observer's activities be very much more simple than they probably actually are in search experiments. In addition to an over-all level of expectancy determined by a nonweighted composite of past events, it is reasonable to suppose that the observer's expectancy fluctuates over short ranges of time in a predictable manner.

Let us assume, for example, that for any given signal the observer's expectancy is exactly determined by the mean intersignal interval. Let us further assume that the observer's ability to estimate time since the last target is free of constant error and shows relatively low variable error. If these assumptions are correct, expectancy should be low immediately after a signal, should increase as the mean intersignal interval is approached, and finally should become quite high as the intersignal interval grows beyond the mean. The mean probability of detection averaged over these periodic fluctuations would correspond to the average intertarget interval. This notion implies that if we look at probability of detection as a function of intersignal interval, we ought to obtain something quite different from the prediction made from the reinforcement hypothesis. Unfortunately, the data thus far available are equivocal, and the equivocality does not only depend upon the confusion caused by the alternate assumptions of occurrence or detection affecting vigilance. There is some evidence from Mackworth (10) and from experiments by the present author (6) indicating that probability of detection increases with increasing intervals between *presentation* of signals; unfortunately, however, both Mackworth's data and the data from the author's experiment are confounded, for purposes of this analysis. One experiment in which spatial location and frequency of intersignal intervals were not confounded with size or intersignal interval obtained negative results (6), both on analysis by stimulus presentation and interval following stimulus detection. A more thorough experimental exploration of the problem is still in order, and it seems safest to regard the issue as unsettled.

Perhaps it should be pointed out that in regard to the question of the temporal distribution of signals, the differences between the two hypotheses we

have been discussing are mainly concerned with one rather elementary point. The question is whether the probability of detection for any given signal is determined exclusively by the temporal location of the signal immediately preceding the signal in question or whether it is determined by a large and rather indeterminate number of signals preceding the signal in question. The assumed activities on the part of the observer differ principally in the nature of reaction to signal feedback. The failure of the reinforcement hypothesis suggests that it is not fruitful to regard the observer as a passive instrument that remembers only the time since the last signal. Rather, it seems more fruitful to regard the observer as a detecting instrument that is continuously performing a kind of averaging of previous input in order to extrapolate the results to future behavior of the search field. This process can be considered to be directly analogous to the important notion of adaptation level, hypothesized by Helson and his associates (e.g., **7**).

Let us briefly consider some problems other than temporal distribution to which the expectancy hypothesis might apply. For example, there is the problem of the spatial distribution of signals. For an experimental hypothesis, we might suppose that the probability of detection for different portions of the field of search could be altered by systematically varying the spatial distribution of signals in the field. In point of fact, it was rather a disappointment to find that no such effect could be experimentally induced (**6**). It could be argued, however, that the experiment conducted by the present author was inadequate; the entire field of search covered only about 30 degrees of visual angle, and the absolute frequency of signals was very high (60/hour), thus minimizing the possibility of differences attributable to spatial distribution of previous signals. Finally, it is possible that the spatial distributions were not extreme enough. Despite all of these arguments as to the inadequacy of the experiment, it seems likely, for radar viewing, that such an effect would be extremely inconsequential in magnitude, if it occurred at all.

The reason for this tentative conclusion stems from an earlier experiment (**4**) from the author's laboratory. In this experiment detectability thresholds (in decibels of attenuation from a fixed signal voltage) were measured under a variety of conditions in which observers were (*a*) instructed to search only a portion of the scope, and (*b*) forced to search only a portion of the scope by having the remainder of the scope masked over. Instructing observers to search only a portion of the scope had no effect upon detectability, but one could improve detectability to the extent of about 2 db. by masking out a portion of the scope (if the masking is done by annuli rather than sectors). From these results it seems unlikely that detectability for PPI-type displays would be very sensitive to changes in spatial distribution of signals, unless the observers were forced not to attend to part of the search field. This fact is probably an inherent property of the viewing of PPI displays, and most likely reflects the effect of a moving sweep and luminance gradient. The result of these factors is that almost inevitably the observer distributes his viewing almost equally over the entire scope. Thus, while some effect of spatial distribution may occur for some visual search tasks, it probably does not occur for all of them.

It is apparent that many problems could arise out of the present statement of the expectancy hypothesis of search behavior. In this brief paper, however, we will look at only one further problem, one that is fundamental to the hy-

pothesis. This problem is provided by the fact that the hypothesis is stated in such a way that the observer's extrapolation of the future behavior of the search field is separated from his vigilance for signals in the field. Such an assumption is necessary if one wishes to avoid the rather artificial implication that vigilance is determined directly by task feedback. It is clear that there are states characteristic of the individual that would not affect his ability to estimate what he *sees the task to be* but which would affect his motivation to participate actively in the task. We may characterize such states as individual differences in vigilance. Having so characterized them, it is our task to discover measures of behavior which predict levels of vigilance expected in some individual on a search task. That such a task is psychometrically feasible is indicated by Jenkins'[4] finding that the within-individual–between-session correlations of probability of detection are high.

The most direct approach to this problem is to look for other measures of behavior in the search situation itself. Among those that have been measured are frequency of false reports (incidence of reporting a signal when none is presented) and reaction time. There are ample data to show that while frequency of false reports varies from one individual to another, it does not correlate with probability of detection (6). On the other hand, Jenkins' analysis of his own data suggests that the reaction time to a pilot signal and probability of detection of the scored signals correlate to a significant degree. This is particularly encouraging, since one would clearly expect vigilance to determine reaction time to a related but unscored signal, and it is entirely reasonable that such a correlation could be independent of the observer's estimate of the future behavior of the task.

It is perhaps a little puzzling that frequency of false report turns out to have no significant correlations (except with rate of signal appearance), but it is at least satisfying (if somewhat *post hoc*) to suggest that expectancy based upon experience swamps any individual differences in tendency to respond to signals. This tentative notion receives some support from the fact that, in the experiments by the present author (6), those individuals who showed a very high false report rate (by verbal report at the close of the experiment) had no higher subjective estimates of the rate of signal appearance than those observers who showed no false reports, or only one or two, during the course of the experiment.

Is an Inhibitory Construct Necessary?

Let us take a brief look at the role of an inhibitory construct in a theory of monitoring or watch-keeping. A general notion of "fatigue" has of course frequently turned up in research on visual tasks. The difficulty of finding a suitable index of fatigue in visual work, however, is ample evidence that the concept, as widely used, is so vague and nonoperational as to be virtually useless. Indeed, Bartley and Chute (1), in their monograph on the subject, showed that they were convinced that the only meaningful referent for the word "fatigue" was to its subjective properties—i.e., to what observers said they felt like.

In the work on visual monitoring, however, a fatigue-like concept has been used in a much more specific and operational context. In Mackworth's studies of visual search he was able to show, under controlled circumstances, that there was a more or less monotonic

[4] See footnote 3.

decline in detectability with time at search over a period of two hours. Various experimental variables introduced by Mackworth convinced him that the decline in performance could be attributed to an active inhibitory process, analogous to the inhibitory processes invented by Pavlov.

Mackworth's notions were elaborated by Broadbent (3) in a paper on classical conditioning and human watchkeeping, and Broadbent pushed the analogy to Pavlov's inhibitory constructs even further. Both Broadbent and Mackworth assume that a continuous decline in detectability is the rule for visual search under monotonous conditions. However, a difficulty with any simple application of the inhibitory construct seems to be that sometimes detectability rises instead of falls with continuous search (6). Sometimes the curves of detectability as a function of time at search fluctuate in a highly irregular manner (6).

It is easy enough to attribute such fluctuations to recovery from inhibition, in the form, perhaps, of disinhibition. But is more difficult to point out just why disinhibition occurs where it does, or, to be more accurate, it is difficult to predict when and where in a given search task disinhibition is apt to occur, unless one introduces special stimulating circumstances to produce it. Thus we must decide, at the least, that the circumstances determining the development of inhibition are very complicated, or that it is easily obscured even in very simple, well-controlled search tasks.

The case against a special inhibitory construct is further bolstered by a consideration of the nature of the disinhibitory stimuli used by Mackworth. In the strict Pavlovian usage, a disinhibiting stimulus is a neutral one; it is hard to know what "neutral" means in the context of visual search, but it is clear that Mackworth's stimuli were meaningful and perhaps motivating to his subjects.

It is also apparent that in visual search an end-spurt can occur if the observer knows something of the duration of the search period, and there is also the suggestion that a warm-up effect occurs, even when the observers are fairly well practiced at the task (6). While there is no compelling reason not to describe all of these effects in terms of the dissipation of inhibition, there is likewise no good reason for doing so. It seems merely cumbersome to propose a state of vigilance, and then to assume that apparent declines in vigilance are due to the active inhibition of the positive process by an additional, negative process.

There is a widespread belief—so widespread, as a matter of fact, that it has almost reached the proportions of folklore—that continuous, monotonous visual search (as in driving on a modern highway) produces a condition akin to hypnotic sleep (whatever that may mean). At any rate it seems clear, because it can be established in the laboratory, that observers may lose vigilance beyond the point of voluntary recovery. That is to say, no matter how hard they try, they cannot increase their attentiveness to the signals being searched for, whereas an extraneous event may restore their ability to observe and detect signals in the visual field.

Despite the fact that at first glance such effects seem most attractively described in terms of inhibition, a more serious glance at these facts, together with certain auxiliary information, suggests that they are more fruitfully conceived of as the absence of some normal, excitatory state. If we turn to the neurophysiological observations on sleep, we see some interesting possibilities that perhaps have a bearing on the problem of vigilance in search. We know, for example, that the waking center in the

hypothalamus, in part at least, must be maintained at a high excitatory level by continued sensory input. Thus the reduction in the sensory input characteristic of monotonous search may reduce the excitability of this or analogous centers in the control of the activity of the organism.

A glance at Kleitman's (8) double theory of sleep shows some interesting comparisons. Kleitman makes a distinction between a wakefulness of necessity and a wakefulness of choice. Wakefulness of necessity is characterized by the polyphasic sleep patterns of many animals and human infants. Whenever such an organism is deprived of sensory input, or the sensory input declines, the activity of the waking center is reduced and the organism sleeps. But to a certain extent the cerebral cortex can control the pattern of wakefulness and sleep; hence the matter of wakefulness of choice. It seems entirely reasonable that the wakefulness of choice depends not only upon the autonomous activity of the cortex, but upon the level of sensory input as well.

This interpretation suggests a good deal more for the theory of vigilance than a mere negation of an inhibitory construct. It suggests the importance of varied sensory input in the voluntary control of vigilance. It is quite possible, for example, that restricting the proprioceptive and tactual input by forcing the observer to remain seated in a rigid position contributes to the decline of vigilance and to an inability of the observer to recover vigilance voluntarily. This is very close to the interpretation given by McFarland, Holway, and Hurvich (9) to their finding that DL's for luminance increase in size as one continues to take measurements; and that, if one allows the observer to stand up and stretch, the DL's recover.

The present view is, therefore, that the problem of maintaining vigilance under conditions of monotonous search can better be viewed as a problem in maintenance of a background sensory input than as the problem of dissipating an accumulation of central inhibition. There are several consequences of this view that are susceptible to experimental test. It suggests that we ought to look for individual differences between observers in measures that reflect an indication of tension level in the individual. Hence, there is the strong suggestion that we look at the interaction between indices of muscle tension, detectability, and signal rate. Furthermore, since there are reasons for suspecting that various personality measures reflect tension levels in the individual, there are reasons for looking at the correlations between personality scores and detectability. The highly promising work of Davis (5), from the Cambridge laboratories, suggests that this is an especially fruitful approach. At any rate, all these problems stem from a notion of vigilance as maintained by excitatory input rather than by inhibition. And, until we have explored these possibilities, it is probably not the best strategy to concentrate exclusively on the properties of an inhibitory construct to account for decline of visual search.

REFERENCES

1. BARTLEY, S. H., & CHUTE, E. *Fatigue and impairment in man.* New York: McGraw-Hill, 1947.
2. BERLYNE, D. E. Attention, perception and behavior theory. *Psychol. Rev.*, 1951, **58**, 137–146.
3. BROADBENT, D. E. Classical conditioning and human watch-keeping. *Psychol. Rev.*, 1953, **60**, 331–339.
4. BUCKLEY, BARBARA B., HANES, R. M., & DEESE, J. Search area and target detectability on a PPI cathode ray tube. *USAF Tech. Report, WADC-TR-503,* 1953.

5. DAVIS, R. Neurotic predisposition and the disorganization observed in experiments with the Cambridge Cockpit. *J. Neurology, Neurosurgery & Psychiatr.*, 1946, **9**, 119–124.
6. DEESE, J., & ORMOND, ELIZABETH. Studies of detectability during continuous visual search. *WADC Tech. Report, WADC-TR-53-8*, 1953.
7. HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297–313.
8. KLEITMAN, N. *Sleep and wakefulness.* Chicago: Univer. of Chicago Press, 1939.
9. MCFARLAND, R. A., HOLWAY, A. N., & HURVICH, L. M. *Studies of visual fatigue.* Cambridge: Harvard Graduate School of Business Administration, 1942.
10. MACKWORTH, N. H. Researches on the measurement of human performance. *Med. Res. Council, Special Report No. 268*, London: His Majesty's Stationery Office, 1950.

(Received January 7, 1955)

Psychological Review
Vol. 62, No. 5, 1955

STATISTICAL THEORY OF DISTRIBUTIONAL PHENOMENA IN LEARNING

W. K. ESTES [1]

Indiana University

One aspect of distribution of practice that recent reviewers (**14**, p. 189; **24**; **25**) seem agreed upon is its persistent refractoriness to any general theoretical interpretation. If *general* is taken to mean *comprehensive*, then I can see no ground for optimism on this point. In all likelihood distributional phenomena are influenced by an extensive set of variables which vary in relative importance from one situation to another. There is no reason, however, why this complexity should be expected to bar progress toward theories which are limited in the sense of dealing with effects of independent variables taken singly or in limited combinations, yet general in the sense of applying to a wide range of situations.

The approach to problems of distribution that will be developed here continues a theoretical program outlined in a recent paper (**6**). The over-all strategy is to start by analyzing in detail the variables of most ubiquitous influence in order to see how much they account for, leaving for later consideration local variables which are important only in restricted types of experimentation.

As in our previous investigation of spontaneous behavioral changes (**6**), preliminary analysis of the distributional problem discloses two independent variables associated with intertrial intervals that probably modify the course of learning under all experimental arrangements. These variables are stimulus fluctuation and interpolated learning. In the preceding study, the former was incorporated into a mathematical model in order to permit derivation of its testable consequences. Since experimental arrangements are essentially the same in studies of spacing as in studies of spontaneous changes, we need only make further application of the same model in order to investigate the role of stimulus fluctuation as a determiner of spacing effects in learning.

It will be convenient to consider separately the two principal aspects of learning distinguished in contemporary theories: (*a*) changes in resistance to extinction or resistance to forgetting, and (*b*) changes in momentary probability, rate, or latency of the response. This distinction will be recognized as corresponding in essentials to that between habit strength and momentary reaction potential in Hull's system (**10**), or that between reflex reserve and reflex strength in Skinner's (**21**). For brevity I shall use the terms *habit strength* and *response strength* to distinguish the two concepts.

At a common-sense level, we should expect that stimulus fluctuation would affect the two types of learning measure differently. To the extent that fluctuation occurs, the make-up of a stimulating situation will vary from trial to trial or from period to period. Therefore, insofar as the effects of this one variable are concerned, rate of learning in terms of change in response strength should be increased by reducing the intervals between trials or periods. On the other hand, resistance to extinction or forgetting should be expected to de-

[1] This paper was prepared during the author's tenure as a faculty research fellow of the Social Science Research Council.

pend on having the response conditioned to as many as possible of the stimuli which might occur during extinction or during a test for retention; therefore, rate of change in habit strength should be directly related to intertrial interval. These notions will be worked out more rigorously in the following pages.

The concepts and assumptions that will be needed for this task have all been presented in a previous paper (6). Only those that are most critical for our present purposes will be reviewed briefly at this point. In order to permit self-selection diets for readers of varying mathematical appetites, I shall again follow the procedure of discussing the model and its applications first at a verbal level and then, in a separate section, in mathematical terms.

CONCEPTS AND ASSUMPTIONS

It is assumed that the stimulus variables associated with any environmental situation undergo constant, random fluctuation. Consequently, of the total population of stimuli available in a given situation over the course of an entire experiment, only a randomly selected subpopulation will be available during any one experimental period. On any one trial, the organism samples the available subpopulation; the probability that any available stimulus element will be sampled on a trial is the parameter θ which has been prominent in other developments of statistical learning theory (7, 19). We shall represent the total stimulus population associated with any experiment by a mathematical set $S_{\bullet}$, the subpopulation available on any one period by a set S, and the subpopulation of temporarily unavailable elements by S'.

There is, then, a continuous, random interchange of elements between S and S'. During a given period, all the elements in the available set S may be conditioned or extinguished while those in S' are unaffected. Since probability of a given response is assumed equal to the proportion of conditioned elements in S, it is possible for response probability to go to unity or zero during a period of conditioning or extinction even though there remain in S' many elements, unconditioned or conditioned, respectively, which may reappear in S during some later period and alter the response probability.

INTERPRETATION OF DISTRIBUTIONAL PHENOMENA

Rate of change in habit strength as a function of spacing. According to our assumptions, there will be available for conditioning or extinction during any one period [2] only a limited portion of the total population S of stimulus elements which may appear in the given experimental situation at any time. Consequently, complete extinction of a response can be achieved only if the series of extinction periods is distributed over a long enough interval so that in the course of random fluctuation every previously conditioned element in the population appears in the available set S during at least one period. And in general, the greater the number of conditioned elements present in both parts of the population $S_{\bullet}$ at the onset of extinction, the longer it will take, in terms of time, trials, or periods, to reach a given criterion of extinction. We have also seen in an earlier paper (6) that, according to the present theory, resistance to forgetting will be directly related to the total density of conditioned elements in $S_{\bullet}$.

Since both of the observable symptoms of habit strength vary directly

[2] To avoid circumlocution, the term *period* will be used here to refer to any of the following: (*a*) a period during which conditioning or extinction is carried to a criterion, (*b*) a period of fixed duration, (*c*) a block of trials, or (*d*) even in some circumstances a single trial.

with total density of conditioned elements, we need only consider the relationship between the latter variable and conditions of spacing. The predicted effects of varying the length of the rest interval between successive conditioning periods follow directly from the sampling model. Rate of increase in the proportion $F(m)$ of conditioned elements in S_n will be maximal if the interval is long enough so that the available situation S during each successive period represents a new random sample from the total population. As the interval is reduced, there is less opportunity for elements that are in S, and thus become conditioned, during one period to escape and be replaced by unconditioned ones before the next period, and consequently rate of growth of $F(m)$ is retarded. These considerations give rise to the family of curves illustrated in the lower panel of Fig. 1, which has interperiod interval as a parameter. Following the first period, steepness of the curve is a uniformly increasing function of the interval. The negative acceleration of these curves is due, of course, to the fact that there will be fewer unconditioned elements remaining to be sampled after each successive period. Analogous considerations hold for changes in $F(m)$ during extinction so that, other things equal, the "resistance to reconditioning" will be an increasing function of the interval between successive extinction periods.

Rate of change in response strength as a function of spacing. Changes in response strength during learning may be followed empirically by obtaining measures of response probability, rate, or latency at the beginning of each successive period. For simplicity, we will consider only probability. It can be shown (5) that all arguments that hold for probability carry over directly to the other measures.

Rate of change in response probability during learning will be less simply related to the interperiod interval than was the rate of change in habit strength. If we compare probability curves for a relatively massed and a relatively spaced group, the former will certainly be expected to show the steeper rise at the outset of conditioning, for between the first and second periods fewer of the newly conditioned elements will escape from S if the interval is short. On the other hand, the gain per period in total number of conditioned elements is greater for the spaced group. From this fact we might guess that as the series goes on, the theoretical curve for the spaced group will overtake and then surpass the curve for the massed group. As a matter of fact this surmise can be proved to be correct, as we shall demonstrate in a later section. The principal characteristics of the family of curves having interperiod interval as a parameter are illustrated in Fig. 2.

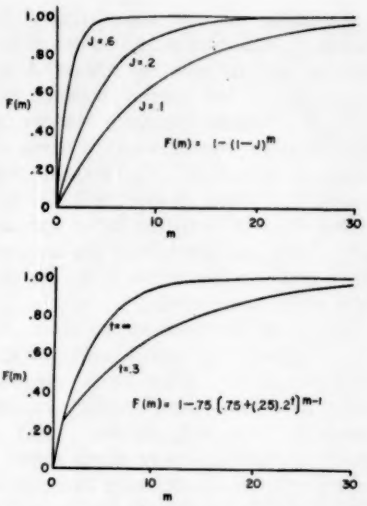

FIG. 1. Proportion of conditioned elements, $F(m)$, in the total population as a function of conditioning periods, m. In the upper panel the parameter is J, the relative size of the available set S during any period; in the lower panel the parameter is t, duration of the interperiod interval.

According to our assumptions, the course of extinction over a series of extinction periods will be influenced by the length of the interval between periods in precisely the same manner as the course of conditioning. Curves of response probability for massed and spaced extinction groups will be expected to cross, with the latter starting out higher but overtaking the massed curve and running below it during the later extinction periods.

So far we have considered only uniform sequences of conditioning or extinction periods, since these are the cases that have been studied experimentally. Suppose, however, that a group of organisms were administered a mixed series of conditioning and extinction periods. Predictions can be generated from the present theory concerning the course of learning over any specified series. An interesting special case arises if we modify the usual partial reinforcement procedure by taking a block of trials or an experimental period rather than a single trial as the unit relative to which reinforcement is randomized. A theoretical account of this case is given in the following section and illustrated in Fig. 2. An experiment which provides a preliminary test of the theoretical predictions has recently been reported (12).

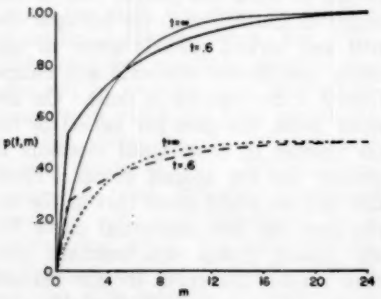

FIG. 2. Response probability in the temporary situation S as a function of number of conditioning or extinction periods, m. The two upper curves represent a sequence of conditioning periods, the two lower curves a random sequence of conditioning and extinction periods with π equal to .5. In each case the parameter is t, the interperiod interval.

MATHEMATICAL DEVELOPMENT OF DISTRIBUTIONAL THEORY

Now it is desirable to show more rigorously how the assumption of random stimulus fluctuation leads to detailed quantitative predictions concerning spacing phenomena. A mathematical model incorporating the fluctuation concept has been presented in a preceding paper (6). Here we shall review briefly a few important terms and relations, and then proceed to apply the same model to problems of spacing.

As before it will facilitate exposition to restrict the formal treatment to highly simplified conditions. We will assume that the behaviors available to the organism in a given situation are categorized into two mutually exclusive and exhaustive classes, A and $\bar{A}$, the former being the response recorded; that each stimulus element in S_* is conditioned either to A or to $\bar{A}$; that probability of response A at any time during an experimental period is equal to the proportion of elements in S conditioned to A; and that all elements in S^* have equal probabilities of appearing in S.

Let t represent duration of the interperiod interval; N and N' the mean numbers of elements in S and S', respectively; C the probability that an element in S' during period $m-1$ appears in S during period m; and B the probability that an element in S during period $m-1$ is still there during period m. Two important relationships derived in the preceding paper (6) are

$$C = J(1 - a^t) \qquad [1]$$

and

$$B = J + (1 - J)a^t, \qquad [2]$$

where J is the ratio of N to $N + N'$, and a is a slope constant depending on rate of stimulus fluctuation.

Now we are ready to derive predictions concerning the course of learning as a function of trial spacing.

Rate of change in habit strength as a function of trial spacing. We require the probability $F(m)$ that an element is conditioned to response A after m learning periods when the interperiod interval is t and all elements present in S during any given period become conditioned to A. The latter assumption implies that the proportion J of the unconditioned elements in S_n will become conditioned during the first period and the proportion C on each subsequent period, yielding the relations

$$\begin{aligned} F(1) &= (1-J)F(0) + J \\ F(2) &= F(1) + C[1 - F(1)] \\ &= (1-J)F(0) + J + C \\ &\quad \times [1 - (1-J)F(0) - J] \\ &= 1 - (1-J)[1 - F(0)] \\ &\quad \times (1-C) \end{aligned}$$

and so on. It can be proved by mathematical induction that in general

$$\begin{aligned} F(m) = 1 &- (1-J) \\ &\times [1 - F(0)](1-C)^{m-1}. \end{aligned}$$

Since from equation 1,

$$C = J(1 - a^t)$$

we have finally

$$\begin{aligned} F(m) = 1 &- (1-J)[1 - F(0)] \\ &\times (1 - J + Ja^t)^{m-1} \quad [3] \end{aligned}$$

and it can be seen that $F(m)$ is an increasing function of the interperiod interval t for all m. This function is illustrated in the lower panel of Fig. 1 with t as a parameter. Under the conditions assumed here, varying the size of the training unit, i.e., the duration of a conditioning period or the number of trials per daily block, will simply result in modifying the value of J, yielding the family of curves illustrated in the upper panel of Fig. 1. Analogous considerations hold for changes in $F(m)$ during extinction.

Comparison of the upper and lower panels of Fig. 1 might lead one to surmise that, beyond a moderate rest interval, the advantage of distribution of practice will come chiefly from dividing a given number of training trials into a larger number of spaced blocks. As a matter of fact this prediction does follow from the theory, not only for the highly restrictive conditions assumed in the derivations of this paper, but quite generally. For any situation in which the learning curve within a daily period is negatively accelerated, the total number of trials needed to attain a criterion of learning will increase with the number of trials per daily period, the function approaching linearity as the number of trials per period becomes large. This relationship is attributable to the fact that fewer elements are conditioned on the second trial of a daily period than would have been conditioned if this trial had been postponed until the next period, and a similar argument holds for each succeeding trial. Once the daily period is made large enough so that nearly all of the available elements are conditioned to the response in question, further increase in the number of trials per period will simply yield a proportionate increase in the total number of trials to a criterion.

Rate of change in response probability as a function of trial spacing. Let us represent by $p(t,m)$ the probability of response A in S at time t following the mth of a series of learning periods, the periods being uniformly spaced with interperiod interval equal to t. First we consider a series of acquisition periods with all elements present in S during a period becoming conditioned to response A. We can write immediately:

$$\begin{aligned} p(t,m) = \frac{1}{N}&[NB + \{(N \\ &+ N')F(m) - N\}C] \end{aligned}$$

where B, C, and $F(m)$ are as defined above. The first term inside the brackets is the number of conditioned elements in S at time t that are held over from the preceding experimental period, while the second term is the number of conditioned elements that were outside S at the end of the preceding period but drifted into S by time t. Simplifying by means of the relation $J = N/N + N'$, and expanding C and B by means of equations 1 and 2, we arrive at the function

$$p(t,m) = B + \left(\frac{F(m)}{J} - 1\right)C$$
$$= 1 - (1 - J)[1 - F(0)](1 - a^t) \times (1 - J + Ja^t)^{m-1} \quad [4]$$

which for all its forbidding appearance has only three free parameters and is relatively simple in form, as illustrated in Fig. 2. It will readily be seen that when m equals 1, response probability will be inversely related to the interperiod interval t. For sufficiently large m the derivative of $p(t,m)$ with respect to t will be positive and therefore response probability will be directly related to t. Therefore, as shown in Fig. 2, any pair of otherwise comparable massed and spaced curves will cross exactly once.

The corresponding function for a series of extinction periods during which all elements present in S are unconditioned from A, i.e., conditioned to an opposed response $\bar{A}$, is obtainable in precisely the same way and turns out to be

$$p(t,m) = F(0)(1 - J) \times (1 - a^t)(1 - J + Ja^t)^{m-1} \quad [5]$$

Random reinforcement. Finally we consider a series of periods run under the same conditions assumed in the derivation of equations 4 and 5 except that acquisition and extinction periods appear in a random order. It can be verified by substitution from equation 4 that if period $m + 1$ is an acquisition period then

$$p(t,m + 1) = (1 - C)p(t,m) + C$$

whereas if it is an extinction period, it follows from equation 5 that

$$p(t,m + 1) = (1 - C)p(t,m).$$

Again C is being used as an abbreviation for $J(1 - a^t)$. If acquisition and extinction periods occur with relative frequency π and $1 - \pi$ respectively, the expected response probability after period $m + 1$ will be the average of the two expressions preceding:

$$E[p(t,m + 1)] = \pi[(1 - C) \times p(t,m) + C] + (1 - \pi) \times (1 - C)p(t,m)$$
$$= (1 - C)p(t,m) + \pi C. \quad [6]$$

Solving this difference equation by standard methods we obtain

$$p(t,m) = \pi - K(1 - C)^m$$

where K is an arbitrary constant, and finally

$$p(t,m) = \pi - (1 - J)[\pi - F(0)] \times (1 - a^t)(1 - J + Ja^t)^{m-1} \quad [7]$$

when K has been eliminated by setting the expression for $p(t,1)$ obtained from equation 7 equal to the mean value of $p(t,1)$ computed from equations 4 and 5. Since the quantity $(1 - J + Ja^t)$ is a fraction between zero and one, the second term on the right side of equation 7 tends to zero as m increases, and response probability approaches π as an asymptote regardless of the value of J or the duration of the interperiod interval. The dashed curves of Fig. 2 illustrate the effect of varying intertrial interval in a series with $\pi = .5$.

It should be noted that the expressions for response probability derived in this paper refer to expected values for individual organisms or for groups of organisms with like values of the parameters: J and a. However, it is not

likely that the errors of approximation involved in applying the derived functions to averaged data from groups of experimental subjects will be large relative to the errors of measurement involved in ordinary learning experiments.

EMPIRICAL RELEVANCE AND ADEQUACY

What I have tried to accomplish in this paper and the one preceding (6) is to work out in systematic fashion the behavioral effects of stimulus fluctuation which is strictly random with respect to experimental variables. To the extent that distributional phenomena prove to be attributable to this source of variation, the burden of explanation upon such constructs as inhibitory potential will be commensurately reduced. To be sure, the conclusions we have reached are not independent of the specific techniques of conceptualization and derivation used, but it seems doubtful that the main features will prove very sensitive to the choice of mathematical techniques.[3]

In applying the theory to learning experiments, we are led to fractionate the empirical material into three main groups: (*a*) experiments involving continuous work or very short rest intervals; (*b*) experiments involving intermediate intervals, ranging from perhaps a minute up to a day or two; and (*c*) experiments involving long intervals. It is in the intermediate range that stimulus fluctuation theory should be expected to play a major role in accounting for experimental phenomena. At very long intervals, or under certain special conditions, such as those of retroactive inhibition experiments, effects of interpolated learning become important; space does not permit discussion of these effects in the present paper. At very short intervals, variables come into play that have not been taken account of in the statistical theory, e.g., response-produced stimulation, physiological changes in effectors; and even if the predicted effects of stimulus fluctuation are present they will be masked by effects of other variables.

Two principal methods are available for assaying the empirical adequacy of the theory. The first is to carry out new experiments especially designed to satisfy the simplifying assumptions made in the theoretical derivations. The second is to review the experimental literature and attempt to provide a unified theoretical account of the already established facts concerning the role of spacing at intermediate intervals.

Only the former method can be entirely free of bias, and consequently publication of the theory has been delayed until several of the indicated experiments could be carried out. One of these was a study of spontaneous recovery of a bar-pressing response in relation to degree of extinction and amount and spacing of reinforcement (8); another was a study of successive acquisitions and extinctions of a jumping habit in relation to schedule of reinforcement (13); a third was a study of successive reversals of a left-right discrimination with varying numbers of trials per daily period (11). Detailed theoretical interpretations of the data of these experiments have been reported elsewhere. It is to researches of this kind that we must look for sharper de-

[3] As a matter of fact, in an earlier version of the present model, the assumptions were expressed in the form of differential equations and the theorems were derived by means of the calculus. Subsequently I found it desirable for reasons of mathematical rigor and elegance to rework the theory utilizing the method of linear difference equations which has been introduced into learning theory by Bush and Mosteller (2), but none of the testable consequences of the theory were altered. The earlier formulation was presented to a Colloquium at the University of Illinois in May, 1950, and to the SSRC Summer Seminar on Mathematics Models for Behavior in 1951, and I have benefited greatly from the discussion and criticism contributed on those occasions.

limitation of the area of application of the theory and for indications as to desirable lines of revision or supplementation.

A comprehensive review of the literature on distributional phenomena cannot be attempted here, but it will be noted that a variety of rather well-established empirical relationships can be given a uniform and coherent interpretation in terms of theoretical relationships derived in this paper. It may be helpful to bring together at this point a brief summary of these relationships together with appropriate references to the experimental literature.

a. Resistance to extinction is directly related to number and spacing of conditioning periods (**8**, **17**, **18**).

b. Retention is increased by spacing of learning periods, the advantage of spacing increasing with the retention interval (**14**, pp. 156–158; **16**, p. 508).

c. Retention is increased by overlearning (**9**; **14**, pp. 377–380).

d. Learning tends to be faster at beginning of series with massed practice, later in series with spaced practice: Conditioning (**1**, **22**); Extinction (**4**, **20**, **23**).

e. Number of trials to asymptote of learning is directly related to number of trials per daily period (**3**, pp. 155–156; **15**, p. 320).

f. Retention is directly related to the proportion of common elements in learning and recall situations (**9**; **14**, pp. 449–450; **19**).

The rather surprising prediction that asymptote of response probability under random reinforcement should be independent of trial spacing has not been listed, since only one relevant empirical study (**12**) is so far available.

All the relationships included have been discussed above except for the last, and this is clearly a direct consequence of the basic assumptions of the theory. In all cases, except perhaps for item *d*, there seems to be unequivocal correspondence of fact and theory. In the case of item *d*, the predicted relationship has been obtained more regularly during extinction than during conditioning. Possibly the difference is attributable to the fact that comparisons of massed and spaced trial groups are better controlled during extinction when such extraneous variables as satiation and habituation of UR's are excluded.

Several additional statements derivable from the theory seem to be almost certainly true, although not yet established by experiment. One of these is the prediction that resistance to extinction will be increased by additional training given after the curve of learning in terms of response probability or latency has reached an asymptote; this relationship seems to have been assumed by Skinner in his formulation of the reflex reserve concept (**21**) but has not, to my knowledge, been subjected to a formal check by experiment. A second example is the prediction that resistance to reconditioning will be directly related to spacing of extinction periods.

The principal set of relevant empirical phenomena not handled at all by the present theory concerns certain properties of response decrement under extremely massed trials. In some experiments with very short intertrial intervals or with conditions of continuous work, as in pursuit rotor performance, asymptotes of intraperiod learning are depressed so that it becomes impossible to achieve the simplifying condition, assumed in the derivations of this paper, that all stimuli available during a period become conditioned to a designated response. In the same experiments reminiscence is commonly observed following rest. These effects apparently are not predictable from the fluctuation theory. The question remains open whether they can be accounted for in terms of variables not so far incorporated into the theory, e.g., response-produced stimulation. For the present these phenomena will provide a source of comfort to those who prefer to terminate analysis at the "molar" level of Hull's (**10**) reactive inhibition construct.

SUMMARY

A mathematical model embodying the concept of random stimulus fluctuation has been found to describe in detail numerous phenomena of spontaneous recovery, regression, or short-term forgetting, and effects of trial spacing at intermediate intervals. This information should be of aid in appraising alternative strategies for theoretical development in this area. It can now be seen that if neo-Pavlovian conceptions of inhibitory potential are to be maintained without major revision, it will be necessary to assume that the amount of random variation in standard experimental situations is negligible. On the other hand, if the amount of random variation is assumed to be appreciable, then a substantial class of distributional phenomena can be accounted for without additional hypotheses.

REFERENCES

1. BARON, M. R. The effect of long intertrial intervals on the limit of eyelid conditioning. *J. exp. Psychol.*, 1952, **44**, 438–441.
2. BUSH, R. R., & MOSTELLER, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, **58**, 313–323.
3. DEESE, J. *The psychology of learning.* New York: McGraw-Hill, 1952.
4. ELLSON, D. G. Successive extinctions of a bar-pressing response in rats. *J. gen. Psychol.*, 1940, **23**, 283–288.
5. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
6. ESTES, W. K. Statistical theory of spontaneous recovery and regression. *Psychol. Rev.*, 1955, **62**, 145–154.
7. ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
8. HOMME, L. E. Spontaneous recovery from extinction in relation to number of reinforcements, spacing of acquisition, and duration of initial extinction period. Unpublished doctor's dissertation, Indiana Univer., 1953.
9. HOVLAND, C. I. Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 613–689.
10. HULL, C. L. *The principles of behavior.* New York: Appleton-Century-Crofts, 1943.
11. LAUER, D. W., & ESTES, W. K. Rate of learning successive discrimination reversals in relation to trial spacing. *Amer. Psychologist*, 1953, **8**, 384. (Abstract)
12. LAUER, D. W., & ESTES, W. K. Observed and predicted terminal distributions of response probability under two conditions of random reinforcement. *Amer. Psychologist*, 1954, **9**, 413. (Abstract)
13. LAUER, D. W., & ESTES, W. K. Successive acquisitions and extinctions of a jumping habit in relation to schedule of reinforcement. *J. comp. physiol. Psychol.*, 1955, **48**, 8–13.
14. MCGEOCH, J. A., & IRION, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.
15. MUNN, N. L. *Handbook of psychological research on the rat.* New York: Houghton Mifflin, 1950.
16. OSGOOD, C. E. *Method and theory in experimental psychology.* New York: Oxford Univer. Press, 1953.
17. REYNOLDS, B. Extinction of trace conditioned responses as a function of the spacing of trials during the acquisition and extinction series. *J. exp. Psychol.*, 1945, **35**, 81–95.
18. ROHRER, J. H. Experimental extinction as a function of the distribution of extinction trials and response strength. *J. exp. Psychol.*, 1947, **37**, 473–493.
19. SCHOEFFLER, M. S. Probability of response to compounds of discriminated stimuli. *J. exp. Psychol.*, 1954, **48**, 323–329.
20. SHEFFIELD, V. F. Resistance to extinction as a function of the distribution of extinction trials. *J. exp. Psychol.*, 1950, **40**, 305–313.
21. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.
22. SPENCE, K. A., & NORRIS, EUGENIA B. Eyelid conditioning as a function of the intertrial interval. *J. exp. Psychol.*, 1950, **40**, 716–720.
23. STANLEY, W. C. Extinction as a function of the spacing of extinction trials. *J. exp. Psychol.*, 1952, **43**, 249–260.
24. TOLMAN, E. C., & POSTMAN, L. Learning. In *Annu. Rev. Psychol.*, 1954, **5**, 27–56.
25. UNDERWOOD, B. J. Studies in distributed practice: XII. Retention following varying degrees of original learning. *J. exp. Psychol.*, 1954, **47**, 294–300.

(Received December 22, 1954)

Psychological Review
Vol. 62, No. 5, 1955

RELATIONSHIPS BETWEEN SHAME AND GUILT IN THE SOCIALIZING PROCESS

DAVID P. AUSUBEL

Bureau of Educational Research, University of Illinois

Guilt is one of the most important psychological mechanisms through which an individual becomes socialized in the ways of his culture. It is also an important instrument for cultural survival since it constitutes a most efficient watchdog within each individual, serving to keep his behavior compatible with the moral values of the society in which he lives. Without the aid rendered by guilt feelings, child rearing would be a difficult matter indeed. If children felt no sense of accountability or moral obligation to curb their hedonistic and irresponsible impulses, to conform to accepted social norms, or to acquire self-control, the socializing process would be slow, arduous, and incomplete. Sheer physical force, threat of pain, deprivation, and punishment, or withholding of love and approval would be the only available methods—combined with constant surveillance—to exact conformity to cultural standards of acceptable behavior. And since it is plainly evident that the maintenance of perpetual vigilance is impractical, that fear alone is never an effective deterrent against antisocial behavior, and that the interests of personal expediency are not always in agreement with prescribed ethical norms, a social order unbuttressed by a sense of moral obligation in its members would enjoy precious little stability.

Within recent years, a number of social anthropologists (3, 7, 8, 9) have advanced the notion that guilt is not universally present or prominent as a sanction in mediating and sustaining the culture. Instead they have identified guilt as a unique property of the characterology of individuals who as children experience the kinds of relationships with parents allowing for "superego" formation, that is, typically of persons growing up in cultures adhering to the Judaic-Christian tradition. Thus, among the Samoans, Iatmul, and Balinese, Margaret Mead believes that the culture is primarily transmitted through such external sanctions as expediency in conforming to the rules (8, p. 514), shared, undifferentiated fear (9, p. 369), and anticipation of physical reprisals (9, p. 370), respectively. Ruth Benedict minimizes the importance of guilt and emphasizes the role of shame in regulating the social behavior of the Japanese (3, p. 222). Leighton and Kluckhohn make the same point in reference to the Navaho: "Sensitivity to shame . . . largely takes the place that remorse and self-punishment have in preventing anti-social conduct in white society" (7, p. 106). And even in characterizing the moral development of adolescents in our own society who undergo peer-group rather than parent-regulated socialization, Margaret Mead comments that "shame, the agony of being found wanting and exposed to the disapproval of others, becomes a more prominent sanction behind conduct than guilt, the fear of not measuring up to the high standard which was represented by the parents" (8, p. 520).

In this paper we shall be concerned with a critical examination of the criteria that these anthropologists have used in differentiating between shame and guilt. We shall attempt to show that although the two kinds of sanc-

tions are distinguishable from one another, they are nevertheless neither dichotomous nor mutually exclusive, and that the development of guilt feelings is not dependent upon highly specific aspects of a unique kind of parent-child relationship. Before turning to this task, however, it might be profitable to undertake a logical analysis of the developmental conditions under which the capacity for acquiring guilt behavior arises, as well as some of the basic relationships between shame and guilt. Because of the paucity of experimental or naturalistic evidence in this area of theoretical inquiry, it should be self-evident that the following paradigm is offered as a system of interrelated hypotheses rather than as an exposition of empirically established facts.

THE DEVELOPMENTAL ORIGINS OF GUILT

Guilt may be conceptualized as a special kind of negative self-evaluation which occurs when an individual acknowledges that his behavior is at variance with a given moral value to which he feels obligated to conform. It is a self-reaction to an injured conscience, if by conscience is meant an abstraction referring to a feeling of obligation to abide by all internalized moral values. The injury consists of a self-perceived violation of this obligation. Hence, in accordance with this formulation, one might hypothesize that before guilt feelings can become operative, the following developmental conditions must apply: (*a*) the individual must accept certain standards of right and wrong or good and bad as his own, (*b*) he must accept the obligation of regulating his behavior to conform to whatever standards he has thus adopted, and must feel accountable for lapses therefrom, and (*c*) he must possess sufficient self-critical ability to recognize when a discrepancy between behavior and internalized values occurs.

It goes without saying that none of these conditions can ever be satisfied at birth; under minimally favorable circumstances, however, all human beings should be potentially capable of acquiring the capacity for guilt behavior. Culture may make a difference in the form which this behavior takes and in the specific kinds of stimuli which instigate it; but the capacity itself should be so basically human and so fundamental to the sanctions by which social norms are maintained and transmitted to the young in any culture that differences among individuals within a culture would probably be as great as or greater than differences among cultures.

It is theoretically possible, of course, that in certain extreme cases a culture may be so anarchic and unstructured in terms of the obligations it engenders in its members that the potentiality for guilt experience is never realized, as, for example, among the Dobu (2, p. 131). Ordinarily, however, we might expect that guilt feelings would be found universally; and, hence, the burden of proof regarding their alleged absence in a given culture more properly rests with the investigator making the allegation.

Despite the probable existence of many important culturally conditioned differences in children's acquisition of guilt behavior, there are presumptive grounds for believing that considerable communality prevails in the general pattern of sequential development. Such communality would be a product of various uniformities regarding (*a*) basic conditions of the parent-child relationship, (*b*) minimal cultural needs for socialization of the child, and (*c*) certain gross trends in cognitive and social growth that prevail from one culture to the next.

The cultural basis of conscience development in the individual may be found in the potent need of both parents and society to inculcate a sense of responsibility in the child. Not only the physical survival of its members, but also the perpetuation of its selective way of life is contingent upon the culture's degree of success in this undertaking. Thus, the attenuation of infantile irresponsibility might be considered part of the necessary process of ego devaluation and maturation that presumably characterizes personality development in all cultures. Socialization demands the learning of self-control and self-discipline, the subordination of personal desires to the needs and wishes of others, the acquisition of skills and self-sufficiency, the curbing of hedonistic and aggressive impulses, and the assimilation of culturally sanctioned patterns of behavior. It seems highly unlikely that any of these propensities could become thoroughly stable before conscience is firmly established.

We might postulate that the first step in the child's development of conscience involves his assimilation of parental values and standards.[1] Having no other frame of reference for judgments of good and bad, the prestige suggestion of parents easily holds sway. But acceptance of these values does not obligate him—in the absence of a still undeveloped sense of moral responsibility—to regulate his own behavior accordingly. Lapses on the part of *other* persons are perceived as "bad," but such judgments have no relevance for similar behavior of his *own* when gratification of his *own* impulses is at stake. At this stage of development his behavior can only be directed into acceptable channels by punishment or by anticipation of punishment in the form of pain, deprivation, ridicule, threatened separation from the parent, etc. Conformity to ethical standards, therefore, is "devoid of moral implications because it is only indicative of submission to authority rather than acceptance of it" (1, p. 122).

Behavior can first be regarded as manifesting moral properties when a sense of obligation is acquired. The central hypothesis of the present formulation is that this development typically takes place in children who are accepted and intrinsically valued by parents, and who thereby acquire a derived or vicarious status in consequence of this acceptance. By the fiat of parental acceptance they are provided with intrinsic feelings of security and adequacy despite their manifest dependency and incompetence to fend for themselves. They accordingly become disposed to accept parental values implicitly and unconditionally out of loyalty to the individuals[1] to whom they owe their status and self-esteem. Among other standards assimilated in this uncritical and subservient fashion "is the feeling of moral responsibility or accountability to conform to standards of behavior which have been given ethical implications. . . . Unlike other values which are ends in themselves, moral responsibility has [the] regulatory function of compelling adherence to internalized norms of behavior" (1, p. 135).

It is reasonable to suppose, however, that the inhibition of unacceptable behavior by feelings of personal loyalty and accountability, by the child's recognition of his parent's moral authority, and by his desire to avoid the self-punishing consequences of guilt develops relatively slowly. The old external sanctions of pain, punishment, threat, and ridicule continue to be ap-

[1] In cultures characterized by an extended family group, or where persons other than parents take major responsibility for child rearing, both value assimilation and the subsequent development of moral obligation have from the very beginning a wider social base.

plied, reinforced now by the more meaningful threat of withdrawal of love and approval. Delaying the consolidation of guilt behavior is the slow growth of both the self-critical faculty (making perception of discrepancies between precept and conduct difficult) and of the ability to generalize principles of right and wrong beyond specific situations.

During preadolescence and adolescence, as the child begins to lose his volitional dependence upon parents and becomes more concerned both with a primary status based upon his own competencies and with equal membership in the social order, the basis of conscience and guilt behavior apparently undergoes significant change. As Piaget has shown, greater experience in interpersonal relationships and in playing differentiated roles in complex social activities makes him more inclined to interpret rules of conduct as functional contrivances designed to facilitate social organization and interaction rather than as sacred and axiomatic givens (12, p. 106). Concomitantly, he develops a notion of moral law based upon principles of equity and embodying a system of reciprocal as opposed to unilateral moral obligations (12, p. 387). The absolutism of his moral standards also tends to be weakened by his greater need for self-enhancement as an individual in his own right —a need that no longer favors unconditional adherence to an uncompromisable set of parental standards, and places a greater premium on the value of expediency. Finally, once the individual acquires this functional concept of moral law, conceives of himself as an independent entity striving for primary status in a social community, and recognizes the reciprocal nature of obligation, he completes the process of transferring "his feeling of moral accountability from parents to the moral authority of society" (1, p. 482).

In heterogeneous cultures, values tend to acquire a wider social base during adolescence. The individual is exposed to a variety of ethical standards and can, within limits, choose between alternative moral systems. And where the peer group plays the major role in adolescent socialization, parents are replaced by peers as the chief interpreters and enforcers of the moral code and, to some extent, as the source of moral authority. But neither set of conditions is essential for the maturational changes in the nature of moral organization that take place during and subsequent to preadolescence (8, p. 519; 10, p. 96). In certain homogeneous cultures also, the source of moral values and authority is referred almost from the beginning to persons outside the immediate family circle (7, p. 51).

ATYPICAL DEVELOPMENT OF GUILT BEHAVIOR

The above description of the original and subsequent development of guilt behavior is hypothesized as the more typical course of moral growth in most children. It presupposes that the latter can achieve derived feelings of status through the medium of a dependent parent-child relationship, and can consequently internalize both values and a sense of moral obligation on the implicit basis of personal loyalty. In all cultures, however, a variable number of parents are psychologically incapable of extending acceptance and intrinsic valuation to their offspring. Thus deprived of the self-esteem derived from the fiat of unconditional parental acceptance, such children are from the very beginning obliged to seek primary status and feelings of adequacy on the basis of their own competencies and performance ability. Accordingly, the basis on which they internalize values and moral obligations might be expected to be correspondingly different.

It would serve no useful purpose here to speculate further on the possible atypical courses that conscience development could take in rejected and extrinsically valued children (1). However, there is less reason to believe that rejecting and extrinsically valuing attitudes characterize *all* parents in a few rare cultures than that they may be found among *some* parents in almost all cultures. Hence, we would hypothesize that the probability of finding some conscienceless individuals in every culture is greater than that of finding conscienceless cultures.

CLASSIFICATION OF SHAME AND GUILT

Generally, shame may be defined as an unpleasant emotional reaction by an individual to an actual or presumed negative judgment of himself by others resulting in self-depreciation vis-a-vis the group. This definition of shame is inclusive of instances in which the persons passing judgment are valued either positively or negatively by the individual being judged (9, p. 367), and of both moral and nonmoral causes for the instigation of such negative judgments. Typical examples of nonmoral shame are embarrassment in committing a breach of propriety or in having one's bodily intimacy exposed to public scrutiny (11, p. 350), and "loss of face" resulting from exposure of ignorance or incompetency. Moral shame, on the other hand, is a reaction to the negative moral judgments of others. It is a prominent sanction in many cultures both before and subsequent to the development of conscience.

Moral shame can, in turn, be divided into two categories—internalized and noninternalized. The latter variety occurs when an individual reacts with self-depreciation to the moral condemnation of others, but does not accept the moral value to which he has failed to conform, e.g., a young child may be shamed by being caught in a lie even if he does not accept the judgment that lying is wrong. Essential for this type of shame is the presence of witnesses to or eventual discovery of the misdeed in question. When the value is internalized, e.g., when the child *really* believes that lying is opprobrious, actual discovery is unnecessary; shame can result merely from presumed or fantasied reproach. Under the influence of such unwitnessed shame, the Ojibway may commit suicide (9, p. 367). According to Margaret Mead, internalized shame only occurs when the parent is the interpreter and enforcer of the sanction (9, p. 366). Observation of closely-knit adolescent peer groups, however, seems to indicate that genuine shame may occur when adolescents perceive or fantasy the moral disapproval of their fellows for offenses against norms established and enforced by the group itself.

The shame associated with guilt may be considered a special case of moral shame. In addition to an internalized moral value, a personal obligation to abide by that value is at stake, e.g., the child not only believes that lying is wrong and feels shame by experiencing or imagining the censure of others, but also feels shame in fantasying the reproach of others for violating a moral obligation. In either case an actual audience is unnecessary, but a presumed or fantasied judgment by others is. However, it is important to emphasize at this point that shame is only one component of guilt, the component involving external judgment and sanction. Guilt also involves other *self-reactions* that are independent of the judgment of others, namely, self-reproach, self-disgust, self-contempt, remorse, lowered self-esteem, anxiety, and various characteristic and subjectively identifiable visceral and vasomotor responses. And

conversely, the shame of guilt is only one of many kinds of shame. Thus, the presence of guilt-related shame does not preclude the simultaneous operation of other forms of moral or nonmoral shame unassociated with guilt; nor does the excitation of any of these kinds of shame (associated or unassociated with guilt) preclude the operation of the various self-judgments characterizing guilt feelings.

Shame relies on external sanctions alone. Guilt relies on both internal and external sanctions. The latter sanctions consist of the presumed judgments of others regarding one's lapses with respect to moral obligations and the resulting self-depreciation vis-a-vis the group, as well as the customary social reprisals associated with the misdemeanors arousing guilt. In addition, feelings of guilt have external reference in that they acknowledge accountability for a moral offense against the group.

In the following section we shall critically appraise (*a*) the views of Margaret Mead and of Leighton and Kluckhohn, who hold that the development of guilt behavior requires a parent-child relationship in which an omniscient parent is the referent for and the source of moral authority, in addition to personally administering and interpeting moral sanctions (**9**, pp. 366, 367); and (*b*) Ruth Benedict's position that guilt is shameless, involves no external sanctions, and is concerned with inner convictions of sin (**3**, pp. 222–223).

ETHNOCENTRIC DISTINCTIONS BETWEEN SHAME AND GUILT

Superego Models of Guilt

The first ethnocentric conception of guilt referred to above equates the capacity for experiencing guilt feelings with neo-Freudian notions of superego formation. According to this conception, guilt feelings cannot arise unless the conditions essential for superego formation are not only present in childhood but are also maintained in subsequent years. Hence, guilt-like behavior in our own or other cultures that does not conform to these conditions is categorized as shame. The two basic conditions that are laid down for the development of guilt are: (*a*) the child must accept the parent or parent surrogate as omniscient or as qualitatively superior to himself in a moral sense, and (*b*) he must accept the parent as the source of moral authority, i.e., as the "referent" in whose name moral behavior is enjoined.

In reference to the former condition Margaret Mead states:

> The adults in the society must think of the child as qualitatively different from themselves, in that the child has not yet attained their moral stature, but is subject to innate impulses which, if permitted unchecked expression, would eventuate in adult character different from and morally inferior to that of the parent (**8**, p. 514).

Leighton and Kluckhohn similarly characterize guilt feelings as representing an individual's acknowledgment that he is "unworthy . . . for not living up to the high standards represented by [his] parents" (**7**, p. 106). In explaining why "Navaho sensitivity to shame takes the place of guilt in our society," they go on to say:

> "Conscience" is related to the belief in an omnipotent God who knows all. . . . In white society the child doesn't see the parent in many life situations. [But] in the circumstances of Navaho life a pose of omnipotence or omniscience on the part of parents would be speedily and almost daily exposed (**7**, p. 106).
>
> Only rarely does one hear the utterances on the part of Navaho parents which are so usual among white parents: "Do it because I say it is right," "do it because I say so," "do it because I am your father and children must obey their parents" (**7**, p. 53).

In the first place, one may legitimately question the assumption that children

do not acquire moral obligations unless a highly authoritarian, paternalistic, and hierarchical parent-child relationship exists, unless the parent sets himself up and is perceived as a qualitatively superior paragon of virtue. As long as the conditions for acquiring derived status are fulfilled, it seems reasonable to suppose that moral obligations are incurred provided that parents or parent surrogates make such expectations unambiguously clear. It undoubtedly increases the prestige suggestion with which the standards and expectations of parents are endowed if the latter individuals are perceived as omniscient; but there is neither convincing logical nor empirical justification for designating this criterion as a *sine qua non* for the acquisition of guilt behavior.

A second difficulty confronting the proposition that qualitative hierarchical distinctions must prevail between preceptor and learner in order for guilt feelings to develop is the suggestive evidence of Piaget, referred to above, showing that as children approach adolescence in our culture they become increasingly concerned with moral obligations based on principles of reciprocity (12, pp. 106, 387). Cooperative and functional relationships with peers engender notions of obligation reflecting mutual respect between equals which gradually displace unilateral feelings of obligation based on implicit acceptance of authority (12, p. 106).

The second condition imposed by adherents of the superego model of guilt behavior, i.e., that parents or parent surrogates must serve as the source of moral authority, runs into similar difficulties posed by developmental shifts in the organization of conscience. It presupposes, in the words of Margaret Mead, that

> . . . teachers, the clergy, judicial officials, etc., partake of the character of the judging parent who metes out reward to the individual child (or, later, adult) who makes a satisfactory approximation to the desired behavior, and punishment to him who does not (8, p. 514).

Such a requirement arbitrarily eliminates the possibility of guilt behavior developing in cultures like the Navaho where the authority for moral sanctions is from the earliest days of socialization referred to the group as a whole (7, p. 51). The same criterion rules out the occurrence of guilt behavior among (*a*) adolescents in our own culture who accept the peer group as their major source of moral authority (see earlier quotation from Margaret Mead), (*b*) rejected and extrinsically valued children who never accept the moral authority of parents, but who may accept rational principles of equity and the moral authority of the group, and (*c*) adults in most cultures, who, in accordance with the developmental changes described above, transfer their allegiances from the moral authority of parents to the moral authority of society. The latter process does not necessarily require the intermediate adolescent step of repudiating adult authority and substituting peer standards. Sufficient for this outcome is a transformation of the basis upon which the authority of elders is accepted, i.e., a change from implicit acceptance of axiomatic truth by a dependent and ideationally subservient individual to a more rational acceptance of functionally necessary norms growing out of interpersonal relations by an independent member of adult society (10, pp. 95–96).

To deny that guilt feelings can arise in the above four situations requires summary dismissal of the evidence of everyday experience with the behavioral content of others. It is admittedly difficult to evaluate the subjective emotional experience of other persons, particularly if they happen to be members of different cultures. Logical inference must always supplement the perceptions

of even the most experienced observer. Nevertheless a decision regarding the meaning of behavior can never be reached on the basis of a priori criteria alone, especially when these lack irrefutable logical or face validity. An emotion is adjudged to be guilt or shame on the basis of its behavioral characteristics, its situational excitants, and its subjective properties, and not because it fails to meet certain arbitrary ethnocentric criteria that apply to a single culture or phase of development. In the next section we shall examine some ethnological materials presented by Leighton and Kluckhohn and by Ruth Benedict as illustrative of shame and show how such behavior can be more defensibly interpreted as guilt.

A conception of conscience or guilt behavior that makes no provision for developmental changes in the underlying personality structure upon which such behavior depends cannot be successfully integrated into any self-consistent theory of personality development. A theoretical framework must be found that is broad enough to encompass the guilt behavior of a child of eight and likewise of an adult of eighteen or eighty. We cannot conclude that eighteen-year-old guilt must be shame because it does not conform to an eight-year-old model of guilt.

Personality development is admittedly characterized by continuity as well as by change. A tendency to accept values and obligations uncritically, on the basis of personal allegiances, still continues in adult life. Similarly, early attitudes of unconditional loyalty to moral obligations tend to persist long beyond the span of childhood. Nevertheless although these continuing trends add much to the stability of conscience, they are largely supplanted by different moral orientations and enjoy only substrate existence in a generally changed gestalt. A precipitate of the experiences of the eight-year-old is represented in the behavior of the eighteen-year-old, but this precipitate has a phenomenologically different meaning and significance than its original form a decade earlier.

Pointing up the logical as well as the empirical untenability of a criterion of guilt that requires a qualitatively superior parent or parent surrogate as the source of moral authority is the discrepancy on this score between the two chief examples of allegedly "shaming cultures," i.e., certain North American Indian tribes and the Japanese. Leighton and Kluckhohn characterize the Navaho as a "shaming culture" in part because sanctions are applied laterally rather than from above, and are referred to group rather than to parental authority (7, pp. 105–106). On the other hand, Ruth Benedict, using other criteria, regards the Japanese as a shaming culture despite the fact that the Japanese family is the ethnological example par excellence of unconditionally accepted, paternalistic (if not arbitrarily imposed) authority based upon hierarchical position (3, pp. 55–57, 264). Obviously, then, this criterion for distinguishing between shame and guilt has little claim to universality.

Guilt as a Shameless, Wholly Internalized Conviction of Sin

The second major ethnocentric (and psychoanalytically oriented) conception of guilt assumes that guilt and shame are mutually incompatible and that genuine guilt must be devoid of shame and all external sanctions. Thus, in cultures where shame plays a more prominent role in behavioral control than in our own culture, ethnologists sharing this conception of guilt have insisted that guilt feelings must *ipso facto* be either nonexistent or of negligible significance. Among the Japanese, according to Gorer, "mockery is the most important

and most effective sanction for obtaining conformity in social life" (4, p. 20). Ruth Benedict espouses the same opinion in stating that "shame has the same place of authority in Japanese ethics that 'a clear conscience,' 'being right with God,' and avoidance of sin have in Western ethics" (3, p. 224). Leighton and Kluckhohn claim that Navahos do not internalize the "standards of their parents and elders," and that shame ("I would feel very uncomfortable if anyone saw me deviating from accepted norms") is the major deterrent to antisocial behavior (7, p. 106). Margaret Mead goes even further by ruling out the occurrence of both guilt and shame in the sanction systems of the Samoans (8, pp. 514–515), Balinese (9, p. 369), and Iatmul (9, pp. 369–370), and insisting that only such external sanctions as fear, expediency, and threat of physical reprisals are operative.

This dichotomization of shame and guilt, of internal and external sanctions, in addition to lacking logical self-consistency, simply does not conform to available naturalistic evidence. Both moral and nonmoral shame may be extremely important sanctions in a given culture, but they preclude the occurrence of neither the public (shame) nor the private (self-judgment) aspects of guilt. A culture so unstable and so anarchic that its members fail to respond to internal sanctions (i.e., feelings of moral obligation) is indeed an ethnological rarity. The inhabitants of Dobu (2) may fit this description, but the case made out for the Samoans, Iatmul, and Balinese (8, 9) is far from convincing; and descriptions of Japanese (3, 6) and Navaho (7) behavior tend to disprove the very contentions they are meant to illustrate. Conversely, a culture in which external sanctions are *not* applied at any stage of moral development is also rare enough to qualify as an ethnological oddity.

The evidence seems clear enough that although Japanese and Navaho individuals are more responsive to shame than we are, much of this shame is really the shame of guilt, i.e., the shame accompanying awareness of violated moral obligations. The predominant pattern of parent-child relationships in both cultures is one of acceptance and intrinsic valuation of the child, leading to the first or implicit stage of moral responsibility based upon personal loyalty (3, Ch. 12; 7, Ch. 1). Navaho children acquire a sense of responsibility between five and eight years of age, learning assiduousness in executing their chores. They learn "that they cannot always indulge themselves and that they have duties toward others" (7, p. 58). Navaho parents are strikingly devoted and attentive to the needs of their children (7, pp. 13–33), and children, in turn, even as adults, are very devoted to their parents and siblings (7, pp. 94–100). The Navaho adult is capable of very responsible conduct in relation to his social group despite the fact that responsibility is "divided and shared" (7, p. 107). If such behavior is not indicative of internalized moral obligation, what other credible explanation is there? Surely shame alone could not account for all of it.

Among the Japanese an even stronger case can be made for the importance of conscience and guilt behavior. Few other cultures lay greater stress upon the sanctity of moral obligations. "The strong according to Japanese verdict are those who disregard personal happiness and fulfill their obligations. . . . A man is weak if he pays attention to his personal desires when they conflict with the code of obligations" (3, pp. 207–208). Obligations to others must be fulfilled even if one dislikes the latter personally (3, p. 124). The Japanese

sense of accountability is especially stringent. "Each man must accept responsibility for the outcome of his acts. He must acknowledge and accept all natural consequences of his weakness, his lack of persistence, his ineffectualness" (3, p. 296).

What is even more impressive is the fact that in contrast to prevailing attitudes in our own culture, the self-discipline, the self-restraint, and any of the personal disadvantages incurred in honoring obligations are not regarded as self-sacrifice. The fulfillment of obligations is taken for granted on the grounds of reciprocity, and calls for no applause, reward, self-pity, or self-righteousness (3, pp. 230–233). The notions of mutual dependence and indebtedness to one's society and one's past lie at the very heart of the Japanese concept of obligation. "One repays one's debts to one's forbears by passing on to one's children the care one oneself received" (3, p. 122).

Despite this overwhelming evidence of conscience and guilt behavior, Ruth Benedict is misled by the dichotomized conception of shame and guilt, and insists upon referring to such behavior as shame on the grounds that guilt is supposedly shameless. But, as already pointed out above, what she calls "shame" is actually the shame component of guilt. La Barre paints essentially the same picture of Japanese character structure as she does and reaches a diametrically opposite conclusion. He asserts that the Japanese superego is exceedingly strong (6, p. 329). Nevertheless, even while minimizing the role of guilt among the Japanese, Ruth Benedict is grudgingly obliged to admit the existence of moral self-judgments unrelated to external sanctions.

> They [the Japanese] are terribly concerned about what other people will think of their behavior, and they are also overcome by guilt when other people know nothing of their misstep. . . . Japanese sometimes react as strongly as any Puritan to a private accumulation of guilt (3, pp. 2–3, 222).

We cannot escape the conclusion, therefore, that both guilt and shame, and internal and external sanctions, can and do exist side by side and mutually reinforce each other. The assertion that "true shame cultures rely on external sanctions for good behavior, not, as true guilt cultures do, on an internalized conviction of sin" (3, p. 323) is unsupported by available evidence. The presence of the stock, the pillory, and the ducking stool in the public market place offers eloquent refutation to the statement that "the early Puritans who settled in the United States tried to base their whole morality on guilt" (3, p. 223). Reinforcing most of the moral sanctions that we customarily assign to the domain of conscience is a parallel set of statutes and group pressures enforced by appropriate public reprisals. Even in cultures where moral obligations are highly internalized, we usually find a policeman on the corner giving a friendly nudge to sluggish consciences or a timely warning to impish consciences pondering a brief vacation from duty.

The regulation of conduct by moral self-judgment is also the *final* step in a long developmental process. In controlling the behavior of children in our own culture, parents may differ from Samoan, Balinese, and Iatmul parents in their propensity for premature sermonizing about right and wrong, filial duty, virtue, etc. But during the first few years of life such preachments are but empty verbalisms to our children, who actually respond as do children in all cultures to the accompanying external sanctions of reward and punishment. Guilt feelings also have external reference by virtue of their shame component which involves a self-deprecatory reaction to the presumed or actual moral censure of others for a transgres-

sion of moral obligation. Finally, the external reference of guilt is apparent in the fact that it is reduced by punishment and confession. It always implies an offense against the group which, therefore, can only be pardoned by group action.[2]

Another source for ethnocentric misinterpretation of Japanese guilt behavior lies in their lack of preoccupation with sin. The idea of sin, however, is in no way indigenous to a conscience founded on moral obligation; it merely tends to be associated with a specific variety of guilt behavior that is influenced by certain religious traditions and notions about the original nature of man. Nevertheless, by equating guilt with "an internalized conviction of sin" (3, p. 223), and conscience with "the avoidance of sin" (3, p. 224), Ruth Benedict finds further "justification" for regarding the Japanese as relatively guiltless.

Japanese neglect of the problem of sin, however, stems from two other sources that can hardly be interpreted as detracting from the genuineness of their expressions of guilt. In the first place, since they do not regard indulgence in physical pleasure as inherently evil—providing that it is properly subordinated to moral obligation—they do not share our notion of "original sin" or our conception of human nature as constantly in need of redemption from inherently sinful desires (3, pp. 191–192). Secondly, since the fulfillment of obligation is taken for granted and is not regarded as self-sacrifice or as a victory of virtue over innately evil propensities, Japanese good conscience is not burdened with the self-righteousness that comes from conquering sin; and, similarly, guilt is not "internalized conviction of sin," but simply awareness of the breakdown of moral obligation.

[2] In private confessionals, the priest serves as the moral representative of the group.

Because the Japanese lay great stress on certain "external" (shame) aspects of self-respect, such as acting with circumspection and paying scrupulous attention to the details of propriety, Western observers are prone to deduce that this precludes their association of the term *self*-respect with a criterion such as "consciously conforming to a worthy standard of conduct" (3, p. 219). However, in the light of the foregoing description of Japanese concepts of moral obligation, we are only justified in concluding that ***in addition*** to sharing the meaning which is commonly placed on "self-respect" in our culture, the Japanese value other character traits (i.e., self-restraint, circumspection, respect for social niceties and for the feelings of others) sufficiently to dignify them by the same term that we usually reserve for moral uprightness.

Lastly, it is important to explain why confession and atonement are such relatively inconspicuous aspects of Japanese guilt behavior.[3] The absence of these characteristics that so frequently accompany the expression of guilt in our culture does not prove, as Ruth Benedict suggests, that genuine guilt is lacking, since otherwise it would be relieved by confession (3, p. 223). The Japanese does not, as alleged, avoid confession because his guilt is really only shame and, hence, would be aggravated rather than eased by being made public. The more credible explanation is that although confession would be guilt reducing it would also be too traumatic in view of the tremendous Japanese sensi-

[3] The practice of atonement is not entirely unknown in Japan. "Boys in later elementary school are sometimes confined to the house for kinshin, 'repentance,' and must occupy themselves with . . . the writing of diaries." Similarly, when ostracized by his schoolmates for offenses, the culprit "has to apologize and make promises before he is readmitted" (3, pp. 273–274).

tivity to shame. Under such circumstances, overwhelming feelings of guilt can be relieved less painfully by suicide.

SUMMARY AND CONCLUSIONS

The capacity for experiencing guilt behavior (i.e., a constellation of negative self-reactions to the perceived violation of moral obligation) is conceived as so basically human and so fundamental to the sanctions by which social norms are maintained and transmitted to the young that under minimally favorable social conditions it should develop in all cultures. The only psychological conditions hypothesized as essential for the development of guilt behavior are (*a*) the acceptance of moral values, (*b*) the internalization of a sense of moral obligation to abide by these values, and (*c*) sufficient self-critical ability to perceive discrepancies between internalized values and actual behavior. The development of conscience typically embodies various sequential changes paralleling shifts in the individual's biosocial status. Ordinarily, greater differences in moral behavior may be anticipated among individuals within a culture than among different cultures.

Shame is a self-deprecatory reaction to the actual or presumed judgments of *others*. It may have moral or nonmoral reference. Guilt feelings always involve a special type of moral shame in addition to other negative *self*-judgments. In addition to the shame of guilt, conscience (guilt behavior) is customarily buttressed by such external sanctions as statutory law, public opinion, and ridicule. Thus, the sanctions of guilt are both external and internal in nature.

Various criteria have been advanced by different ethnologists for distinguishing between shame and guilt. It has been alleged (*a*) that guilt behavior can only arise when a hierarchically superior parent or parent surrogate serves as the source of moral authority, (*b*) that genuine guilt feelings can only exist when shame and other external sanctions are not operative, and (*c*) that, subjectively, guilt must be characterized by conviction of sin and need for atonement. Examination of Navaho and Japanese ethnological materials presented in support of these criteria shows that the latter lack both logical and empirical validity. Highly suggestive evidence of the operation of strong moral obligations and of guilt behavior was found in these supposedly prime examples of "shaming cultures."

The problem of shame and guilt is illustrative of a major methodological hazard in the cross-cultural investigation of personality traits, namely, of the ethnocentric tendency to define a given trait in terms of its *specific* attributes in one's own culture. This leads to a perversion of cultural relativism. Instead of demonstrating how a basic human capacity (e.g., guilt behavior) occurs under different conditions and assumes different forms in different cultures, this approach inevitably "discovers" that this very capacity is absent in other cultures because it does not conform to a specific set of ethnocentric crtieria. The generalization is then made that the capacity is nonuniversal, culture-bound, and, hence, not basic. Cross-cultural investigations of personality must, therefore, start with definitions of traits that are both psychologically meaningful and general enough to encompass specific cultural variants.

It is also important to avoid the error of circularity—setting up *a priori conditions* for the emergence of a trait, finding these conditions absent in a given culture, and concluding "empirically" therefrom that the *trait itself* is absent. This error can be avoided in

part by paying attention to the subjective and expressive content of behavior and not merely to the postulated determinants of the trait in question.

REFERENCES

1. AUSUBEL, D. P. *Ego development and the personality disorders.* New York: Grune & Stratton, 1952.
2. BENEDICT, RUTH. *Patterns of culture.* Boston: Houghton Mifflin, 1934.
3. BENEDICT, RUTH. *The chrysanthemum and the sword.* Boston: Houghton Mifflin, 1946.
4. GORER, G. *Japanese character structure.* New York: Institute for Intercultural Studies, 1942.
5. JENKINS, R. L. Guilt feelings—their function and dysfunction. In M. L. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950.
6. LA BARRE, W. Some observations on character structure in the Orient: the Japanese. *Psychiatry,* 1948, **8**, 319–342.
7. LEIGHTON, DOROTHEA, & KLUCKHOHN, C. *Children of the people: the Navaho individual and his development.* Cambridge, Mass.: Harvard Univer. Press, 1947.
8. MEAD, MARGARET. Social change and cultural surrogates. In C. Kluckhohn & H. A. Murray (Eds.), *Personality in nature, society and culture.* New York: Knopf, 1949.
9. MEAD, MARGARET. Some anthropological considerations concerning guilt. In M. L. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950.
10. NICHOLS, C. A. *Moral education among North American Indians.* New York: Teachers Coll., Columbia Univer., 1930.
11. NUTTIN, J. Intimacy and shame in the dynamic structure of personality. In M. L. Reymert (Ed.), *Feelings and emotions,* New York: McGraw-Hill, 1950.
12. PIAGET, J. *The moral judgment of the child.* Glencoe, Ill.: Free Press, 1932.

(Received January 17, 1955)

Psychological Review
Vol. 62, No. 5, 1955

NOTES TOWARD A MOTOR THEORY OF VISUAL EGOCENTRIC LOCALIZATION [1]

JAN H. BRUELL AND GEORGE W. ALBEE

Western Reserve University and Highland View Hospital, Cleveland, Ohio

The authors are engaged currently in a series of studies of space perception in hemiplegic patients. The known relationship between space perception and tonic processes provided the rationale for our studies with hemiplegics whose main problem, by definition, is a disturbance of motor processes. In the process of collecting and analyzing our data we have found that a variety of phenomena of space and movement perception can be ordered, using a limited number of hypotheses.

Space perception refers to impressions of the positions of objects relative to each other and relative to the observer. The convention in present-day treatment of the problem is to refer to judgments of the relative position of objects to each other as *relative localization* and to judgments concerning the position of objects relative to the observer as *egocentric localization.* The present paper will be restricted to a consideration of egocentric localization, and more specifically to the visual perception of direction and egocentric apparent movement.

Visual Egocentric Localization and the Innervation of the External Eye Muscles

Retinal Local Signs and Voluntary Eye Movements

The perception of the egocentric position of visual objects is often attributed to a hypothetical function of retinal receptors. Each retinal point is assumed to have a definite space value or local sign. Conventionally (14) receptors to the right of the median meridian of the retina are given negative space values, and those to the left of the meridian positive space values; receptors in the fovea and median meridian are assumed to have a zero space value, as shown in Fig. 1*a*. In this and succeeding figures it is assumed that the stimulus objects are placed at least six feet from the observer. At this distance convergence is negligible and the binocular situation can be represented by a "cyclopean eye." Objects A and B, which stimulate receptors with negative space values, are perceived as being to the observer's left; objects D and E, stimulating receptors with positive space values, are perceived as being to the right; and object C, projected on the fovea (F), is perceived as being straight ahead.

Obviously this applies only when the observer looks at a point straight ahead, as in Fig. 1*a*. In Fig. 1*b*, if the eye is turned *voluntarily* from C to E, then object C, placed in the objective median plane of the observer, stimulates a peripheral part of the retina. Yet the observer continues to perceive C as being straight ahead of him. That is, a peripheral point on the retina has assumed a space value of zero. On the other hand, object E, placed to the right of the objective median plane, now stimulates the fovea but is perceived as being to the right. Clearly, with voluntary eye-turning, retinal points change their space values. With voluntary eye-turning to the right (Fig. 1*b*), the fovea

[1] This paper will ultimately be part of a larger study of space perception made possible by a grant-in-aid of research from the Cleveland Foundation.

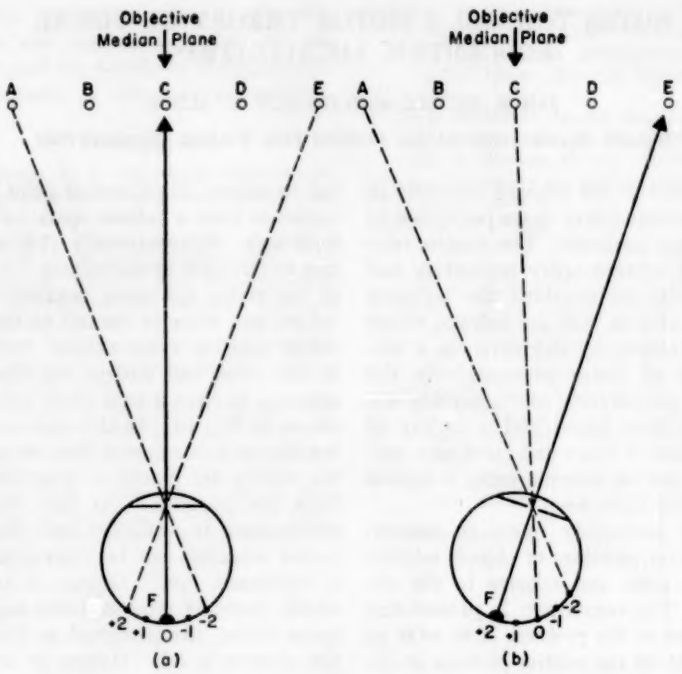

FIG. 1. Schema illustrating concept of retinal space values: (*a*) in primary position of the eye, (*b*) with the eye turned voluntarily to the right.

assumes a positive space value and all other retinal points change their space values correspondingly.

This common observation leads to our first hypothesis: whenever the dextro- or levorotators of the eyes receive *voluntary* innervation, the fovea assumes a positive or a negative space value, respectively, and concomitant changes occur in the space values of all retinal points (Hypothesis 1).

Retinal Local Signs and Involuntary Movements

Figure 2 represents a situation in which the eyes of an observer remain stationary while an object C is moved to the left. Object C stimulates retinal points with space values of -1, -2, -3, etc., in succession. Under these conditions C is perceived as moving to one's left, as it actually does. The same retinal stimulation occurs when an object, C, remains stationary while the eyes are turned to the right. However, whether such retinal stimulation is experienced as movement of C to the left depends upon the voluntary or reflex nature of the eye movement.

As was shown above, in the case of *voluntary* eye-turning, retinal receptors change their space values; and C, which stimulates more and more peripheral points on the retina, continues to be perceived as straight ahead. On the other hand, if the eyes turn *involuntarily* to the right, the stationary object C will appear to move to the left. If, for instance, a subject is rotated coun-

terclockwise around his vertical axis, his eyes will move involuntarily to the right (slow phase of vestibular nystagmus). Object C, placed straight ahead of the subject and rotated with him, will stimulate increasingly peripheral points on the right half of the retina of the involuntarily turning eye. These points have negative space values when the eye is in its primary position. From the fact that under the described conditions the object is seen moving to the left, we conclude that with involuntary eye-turning, retinal points do not change their space values.

Thus we arrive at a hypothesis supplementary to the hypothesis derived from our discussion of the effects of voluntary eye-turning on retinal local signs. There we hypothesized that whenever eye muscles receive voluntary innervation, all retinal receptors, including the fovea, change their space values. The fovea assumes a positive space value if one attempts to turn the eyes to the right, a negative space value if the voluntary effort is toward turning the eyes to the left. Now we hypothesize: when eye position is determined only by *involuntary* innervation of the eye turners, retinal receptors do not change their space values (Hypothesis *2*). For instance, the fovea retains its space value of zero, and any object projected upon the fovea appears to be straight ahead.

These hypotheses, if valid, should permit us to predict the apparent egocentric position of objects from a knowledge of the pattern of oculomotor innervation obtaining in a given situation. We shall now examine certain factors determining the innervation of the eye turners.

Some Factors Determining Eye Position

The position of the eye in its orbit is determined by the innervation, and consequently by the state of contraction

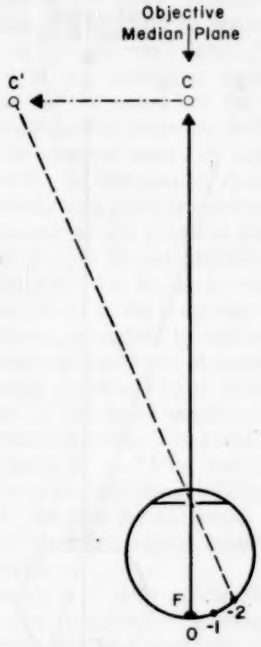

FIG. 2. Schema illustrating retinal stimulation by object moving to the left.

of the six external eye muscles. This innervation may be either voluntary or reflex. The reflex innervation of the eye turners can be brought about by retinal stimulation or it can have its origin in the motor system.

Stimulation of receptors in the periphery of the retina elicits the fixation reflex. The eye turns reflexly in the direction of the stimulus. The reflex eye movement ceases when the stimulus projects on the fovea. Figure *3b* represents an instance of unilateral retinal stimulation. Stimuli are placed only to the right of fixation, causing reflex eye-turning to the right. As the eye turns, retinal stimulation becomes bilateral as receptors on both sides of the fovea are stimulated. Such bilateral stimulation results in reflex innervation of both the levo- and the dextrorotator muscles.

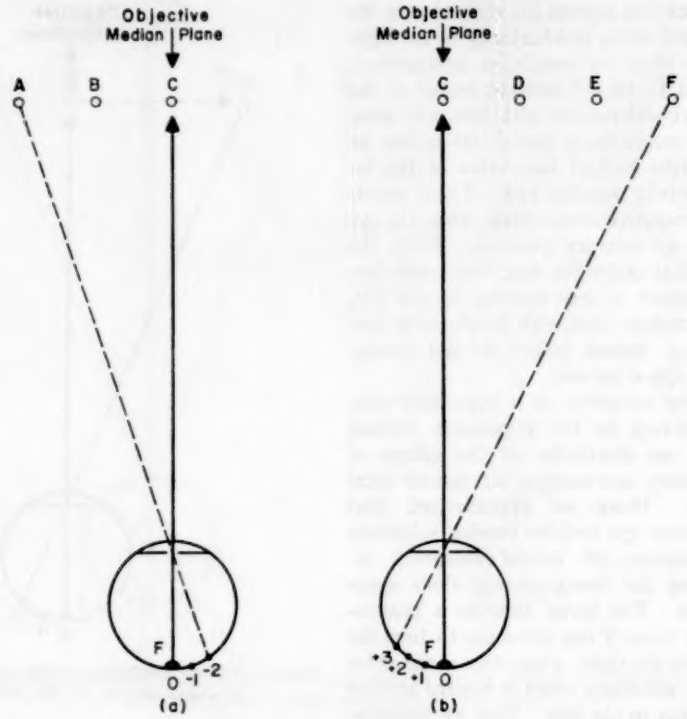

FIG. 3. Schema of visual field extending (*a*) to the left and (*b*) to the right of fixation.

Initially retinal stimulation is asymmetrical, causing stronger innervation of the dextrorotator. When the eye fixates the center of the stimulus field, the fixation reflex will cause innervation of the dextro- and the levorotator, and the eye will cease turning to the right (3).

This analysis assumes that eye position is determined by retinal stimulation only. Nonvisual muscular factors, however, co-determine eye position. In the absence of visual stimulation the eyes tend to assume their "primary position," that is, they tend to look straight ahead. Although it is certain that several reflex mechanisms contribute to this effect, a myotatic reaction is probably involved. When, for instance, the eyes are turned to the right, the levorotators are stretched. This stretching stimulus is thought to elicit a myotatic reflex in which the stretched levorotators contract, thus returning the eyes to their primary position. The recent discovery of kinesthetic receptors in the eye muscles (20, 23) makes the existence of such a myotatic reflex quite plausible. However, as will be seen, whether the eyes are maintained in the median position by this particular reflex or by another reflex mechanism is not crucial to our formulation. In the situation represented in Fig. 3*b*, the fixation reflex acting alone would cause the eyes to turn to the right until the center of the stimulus field is fixated. The myotatic reflex acting upon the

stretched levoturners, however, will tend to arrest the turning of the eyes to the right before the center of the stimulus field is fixated.

It is important to note that if the observer is instructed to fixate continuously object C (Fig. 3*b*), he can do so only by counteracting the reflex tendency of the eyes to turn to the right. He must innervate voluntarily the levorotators.

In short, if the position of the eyes is known for a given stimulus situation, it is possible to infer whether eye position is being maintained by reflex innervation of the eye turners alone, or whether it is co-determined by voluntary innervation. Analyses such as these, in conjunction with our hypotheses concerning the effects of voluntary and reflex innervation of eye turners on retinal local signs, will form the basis for our discussion of a variety of phenomena of egocentric localization.

THE PERCEPTION OF THE MAIN COORDINATES OF EGOCENTRIC SPACE: APPARENT STRAIGHT-AHEAD AND APPARENT EYE LEVEL

The Dietzel-Roelofs Phenomenon

Dietzel (5) and, independently, Roelofs (19) performed experiments which fit exactly the paradigm represented in Fig. 3. A luminous figure, the only visible object in a darkroom, was placed in such a position that one of its edges was in the objective median plane of the observer. The observer had the task of judging the egocentric position of C (Fig 3*b*). Our hypotheses permit us to predict the outcome of such an experiment.

It was shown in the preceding section that continuous fixation of C would be possible only if the levorotators received voluntary innervation. Thus, from our hypothesis concerning the effects of voluntary innervation on the local sign of the fovea (Hypothesis 1) it follows that C would be perceived as placed to the observer's left. The experience of the egocentric position of C would be the same, but the underlying physiological processes would differ if the observer were not instructed to fixate C. In this case his eyes would turn reflexly to a point somewhere short of the center of the stimulus field, say to point D. At D, eye position would be maintained, due to the balanced reflex innervation of the levo- and dextrorotators and, according to our hypothesis concerning the effects of involuntary innervation on retinal local signs (Hypothesis 2), the fovea would retain its space value of zero. The foveally projected point D would be seen by the observer as straight ahead of him, while point C, *relative to D*, would appear to be to his left (relative localization; see introduction).

Similarly, if C were at the right extreme of a stimulus configuration, and if it were placed objectively straight ahead of an observer (Fig. 3*a*), it would appear to him to be to his right.

Dietzel (5) and Roelofs (19) actually obtained these results, and they have been confirmed in numerous recent experiments (3, 18, 21). The general findings of these experiments can be accounted for by our hypotheses. The apparent straight-ahead always tends to shift in the direction of the center of the stimulus field. If this center deviates from what is objectively straight ahead of the observer (Fig. 3), objective and subjective straight-ahead do not coincide.

Ordinarily we are fairly accurate in judging what is straight ahead. That is, objective and subjective straight-ahead *do* coincide. This can be accounted for in the same terms as the described discrepancies between objective and subjective space. Usually the whole visual field is filled with stimuli,

and thus if a point in the objective median plane is fixated, retinal stimulation tends to be bilateral and symmetrical. This results in equal reflex innervation of the levo- and dextrorotators, and eye position is maintained by reflex innervation alone. Under these conditions, according to our second hypothesis, the fovea will have a space value of zero, and the centrally located fixated object will appear to be straight ahead.

A situation outside of the laboratory in which the center of the stimulus field deviates from what is objectively straight ahead occurs during night driving. In Fig. 3*a*, point C might be one of the headlights of an oncoming car. This headlight will be perceptually displaced in the direction of the observer. It is probably a benign illusion as it makes him steer his own car farther away from the passing car than would be required by the objective situation.

Egocentric Localization in Hemianoptics

In the situation diagrammed in Fig. 3*b*, a place near the center of the stimulus field appears to be straight ahead. This deviation of what is apparently straight ahead from what is objectively straight ahead can be produced in anyone in a darkroom. In this situation we know that if the subject fixates C (Fig. 3*b*), only his left occipital lobe will receive stimulation. The situation is comparable to what happens in a hemianoptic where there is a loss of the left half of the visual field. In the darkroom experiment only the left occipital lobe *receives* stimulation, owing to the experimental arrangement. In the hemianoptic with a loss of the left half of the visual field, only the left occipital lobe *responds* to stimulation, because of the injury in the other half of the visual system. Thus one would expect that in such patients apparent straight-ahead would deviate from objective straight-ahead in their daily perception. This expectation is borne out by clinical observation (2) and experiments (7).

The observation that egocentric localization is disturbed in hemianoptics might lead to the conclusion that the perception of the main space coordinates is linked to purely sensory (occipital) visual processes. Therefore, it should be stressed that our interpretation, which relates phenomena of egocentric localization to processes of oculomotor innervation, is in itself in no way incompatible with the observations made in hemianoptics. On the contrary, as we believe that visual egocentric localization is partially determined by optically induced reflex movements of the eye, and as we know that in man the motor branch of such reflexes has its origin in the occipital lobes (15), we would, on these grounds alone, expect egocentric localization to be disturbed in hemianoptics.

The Robert MacDougall Phenomenon

Our discussion so far has centered around phenomena of egocentric localization in the left-right dimension. What applies to the perception of mediality applies also to the subjective horizon. In fact, it was in the up-down dimension that displacements of objects analogous to those observed by Dietzel and by Roelofs were studied first. Robert MacDougall (16) found at Harvard in 1903 that the fixated upper edge of a stimulus figure, if placed at the objective eye level of an observer, appeared to be above eye level. MacDougall's finding has been confirmed in a number of experiments (10, 11), and can be explained in the same terms as the results of Dietzel and of Roelofs. Fixation of the upper edge of a stimulus figure can be maintained only by the

voluntary counteraction of the reflex tendency of the eyes to turn downward toward the center of the stimulus (fixation reflex). The elevator muscles of the eyes have to receive voluntary innervation. This changes the local sign of the fovea from eye-levelness to upness, and the fixated upper edge of the configuration appears to be above eye level. MacDougall's darkroom experiment was inspired by the fact that, in spite of the curvature of the earth, an observer standing on the deck of a ship will perceive the horizon as being above eye level.

APPARENT EGOCENTRIC MOVEMENT

The preceding discussion dealt with the perception of the egocentric position of objectively and phenomenally stationary objects. In what follows we will discuss phenomena of apparent egocentric movement, i.e., we will be concerned with apparent changes in the egocentric position of stationary objects.

If an observer fixates a stationary light C (Fig. 3*b*) and if another light, D, is turned on, he will have a tendency to turn his eyes toward the peripheral light D. In order to counteract this tendency he must voluntarily innervate his levorotators. When light E is turned on, a further increase in the voluntary innervation of the levorotators becomes necessary. The same effects occur when lights F and G are turned on in succession. The increasingly strong reflex innervation of the dextrorotators due to stimulation of more and more peripheral parts of the retina (3) must be counteracted by increasing correspondingly the voluntary innervation of the levorotators.

From our first hypothesis it can be predicted that such a gradual increase in the voluntary innervation of the levorotators would change the space value of the fovea from 0 to -1, to -2, to -3, etc. (cf. above, p. 392). The fixated point C and all other points of the stimulus configuration would appear to move to the left. This prediction is borne out by experiments which we have conducted and which will be published elsewhere.

It is important to note that in this latter situation both object and eye remain stationary. According to the foregoing analysis, apparent movement of the stimulus field will occur if voluntary innervation of one set of eye turners (e.g., the levorotators) increases gradually but does *not* result in actual eye movements. The apparent movement of the stimulus field under such conditions occurs in the direction of the voluntary innervation. Many examples of apparent movement which can be explained in these terms have been described in the literature. We have selected a few examples for presentation here.

Vestibular Stimulation

In developing our hypothesis concerning the effects of involuntary eye movements on retinal local signs, we have already referred to vestibular stimulation. We discussed a situation in which a stationary object stimulates successive retinal points of the eye which is turning as a result of involuntary innervation of the eye turners. In what follows we are concerned with a situation in which, in spite of vestibular stimulation, the eyes remain stationary.

As was already described, vestibular nystagmus of the eyes can be produced by rotating subjects around their vertical axes. With counterclockwise acceleration the eyes will move slowly clockwise (slow phase of nystagmus). If, under these conditions, an observer wants to keep fixating an object placed in front of him and rotating with him, he must counteract the increasing re-

flex innervation of his dextrorotators by increasing the voluntary innervation of the levorotators. According to our first hypothesis and its elaboration above, we would predict that such a progressive increase in voluntary innervation of the levorotators would lead to a gradual change in the space value of the fovea. The fixated object would appear to be moving to the left. Morant (18) has reported experiments which bear out this prediction.

Paralysis of the Eye Turners

The clinical literature describes many cases of peripheral paralyses of the eye turners. Patients with such paralyses experience apparent movement of the stimulus field if they try to turn the paralyzed eye. Helmholtz (9) describes a case of a patient whose dextrorotator of the right eye was paralyzed. Observations were made with the paralyzed eye only. The patient fixated an object C (Fig. 3). If he tried to turn his eye to the right, voluntary impulses sent to the dextrorotator did not result in eye rotation. The patient kept fixating C. With the attempted innervation of the dextrorotator the whole stimulus field appeared to the patient to move to the right. Mach (17) has performed an experiment in which he "paralyzed" the eye by mechanical means. He placed putty around the eye in such a way as to immobilize it. If he attempted to turn the eye to the right, he could not do so. Stimuli projected on his retina continued to stimulate the same retinal points. The voluntary innervation of the dextrorotators, however, changed the space values of all retinal points, and the environment appeared to move to the right.

The apparent movement of the environment observed in patients with eye-muscle paralysis and in Mach's experiment occurs in the direction of the intended eye movement. These phenomena were to be predicted from our hypothesis concerning the effects of voluntary innervation of eye turners on retinal local signs.

Autokinetic Movement

Few phenomena have received as much attention from psychologists as the apparent movement of a so-called autokinetic light. As we will deal elsewhere extensively with this special case of apparent egocentric movement, we limit our present treatment of the subject to one short comment.

The erratic course of the stationary autokinetic light has often been ascribed to involuntary eye movements. This theory had to be abandoned when Guilford and Dallenbach (8) photographed the eye movements occurring during the phenomenon and found them to be slight and not of such a nature as to account for the experience. This finding was interpreted subsequently to mean that oculomotor processes have nothing to do with the autokinetic phenomenon (1). Such a conclusion does not appear to be warranted by the evidence. On the contrary, we have just discussed phenomena in which the absence of observable eye movements in the presence of voluntary innervation of eye turners was an explicit condition for apparent movement. We believe that the apparent movement of the autokinetic light belongs in this same class of phenomena.

Induced Movement

If an observer fixates a light, C (Fig. 3*b*), placed straight ahead of him, and if another light, D, is placed to the right of C, the dextrorotators of the eyes receive reflex innervation. This innervation increases as light D is slowly moved to the right. Fixation of C can be

maintained only by an increase in the voluntary innervation of the levorotators. According to our first hypothesis, such gradual increase in voluntary innervation of the levorotators leads to a progressive change in the space value of the fovea, and it can be predicted that the fixated object will appear to move to the left. Experimental studies (6) bear out this prediction.

The apparent movement of C is induced by the objective movement of D. On the other hand, if C were moved slowly to the left while a peripheral point, D, was held stationary, apparent movement would be induced in D. As fixation of a slowly moving object is maintained by reflex action (pursuit movement), the space values of the retina would not change (Hypothesis 2). As the eye moves and D remains stationary, D would stimulate more and more peripheral points on the left half of the retina (Fig. 2), thereby appearing to move to the right.

Many instances of such induced movement are known to everyone. The most frequently mentioned example is the apparent movement of the moon seen through drifting clouds. Another example is the familiar illusion experienced when one is sitting in a train and the train on the next track slowly starts to move.

SUMMARY

The relationship between phenomena of visual egocentric localization and processes of involuntary and voluntary oculomotor innervation has been pointed out by many earlier investigators (**4, 9, 12, 13, 17, 19**). This paper has attempted to describe the nature of this relationship more precisely. From an examination of empirical data relating to visual egocentric localization, two hypotheses were derived to account for the perception of direction and for some phenomena of egocentric apparent movement. Hypothesis 1 states that when dextro- or levorotators of the eye receive voluntary innervation, the space values for points on the retina change in relation to the direction and degree of innervation. Corresponding changes of retinal space values occur with voluntary innervation of the elevators and depressors of the eye. Hypothesis 2 states that when eye position is determined by reflex innervation alone, retinal points do not change their space values. Factors responsible for voluntary and reflex innervation of eye turners were examined, permitting the knowledge of patterns of innervation under various stimulus conditions. Predictions from the hypotheses concerning the effects of voluntary and reflex innervation on space perception were compared with empirical data, and were found tenable for a variety of perceptual experiences.

REFERENCES

1. BARTLEY, S. H. *Vision, a study of its basis.* New York: Van Nostrand, 1941.
2. BEST, F. Die Bedeutung der Hemianopsie fuer die Untersuchung des optischen Raumsinnes. *Arch. f. d. ges. Physiol.*, 1910, **136**, 248–262.
3. BRUELL, J. H. Visual egocentric localization, an experimental study. Unpublished doctor's dissertation, Clark Univer., 1953.
4. CARR, H. A. *An introduction to space perception.* New York: Longmans, Green, 1935.
5. DIETZEL, H. Untersuchungen ueber die optische Lokalisation der Mediane. *Z. f. Biol.*, 1924, **80**, 289–316.
6. DUNCKER, K. Ueber induzierte Bewegung. *Psychol. Forsch.*, 1929, **12**, 180–259.
7. FUCHS, W. Untersuchungen ueber das Sehen der Hemianoptiker und Hemiamblyopiker. *Z. f. Psychol., 2. Abteilung*, 1920, **84**, 76–169.
8. GUILFORD, J. P., & DALLENBACH, K. M. A study of the autokinetic sensation. *Amer. J. Psychol.*, 1928, **40**, 83–91.
9. HELMHOLTZ, H. *Handbuch der Physiologischen Optik*, dritte Auflage, ergaenzt und herausgegeben in Gemeinschaft mit Prof. Dr. A. Gullstrand, Prof. Dr. J.

von Kries, Prof. Dr. W. Nagel. Leipzig: Leopold Voss, 1910.

10. JAFFE, K. M. Effect of asymmetrical position and directional dynamics of configurations on the visual perception of the horizon. Unpublished master's thesis, Clark Univer., 1952.
11. KADEN, S. E. Effect of directional dynamics of configurations and visually perceived words on the position of the apparent horizon. Unpublished master's thesis, Clark Univer., 1953.
12. KLEINT, H. Versuche ueber die Wahrnehmung. *Z. f. Psychol.*, 1937, **141**, 9–44.
13. KLEINT, H. Versuche ueber die Wahrnehmung. *Z. f. Psychol.*, 1938, **142**, 259–290.
14. KOFFKA, K. *Principles of gestalt psychology*. New York: Harcourt, Brace, 1935.
15. LINKSZ, A. *Physiology of the eye*. Vol. 2: *Vision*. New York: Grune & Stratton, 1952.
16. MACDOUGALL, R. The subjective horizon. *Psychol. Rev. Monogr. Suppl.*, 1903, **4**. Pp. 145.
17. MACH, E. *The analysis of sensation*. Chicago: Open Court, 1914.
18. MORANT, R. B. Factors influencing the apparent median plane under conditions of labyrinthian stimulation. Unpublished doctor's dissertation, Clark Univer., 1952.
19. ROELOFS, C. O. Optische Lokalisation. *Arch. Augenheilk*, 1935, **109**, 395–415.
20. SCOBEE, R. G. *The oculorotary muscles*. St. Louis: C. V. Mosby, 1952.
21. WAPNER, S., WERNER, H., BRUELL, J. H., & GOLDSTEIN, A. G. Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane. *J. exp. Psychol.*, 1953, **46**, 300–307.
22. WERNER, H., & WAPNER, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324–338.
23. WOLTER, J. R. Morphology of the sensory nerve apparatus in striated muscles of the human eye. *Arch. ophthalmol.*, 1955, **53**, 2, 201–207.

(Received for early publication April 1, 1955)

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by JAMES DREVER

Vol. XLVI. Part 2 May 1955 $3.50 net.

HARRY KAY. Learning and retaining verbal material.

PETER H. VENABLES. Changes in motor response with increase and decrease in task difficulty in normal industrial, and psychiatric patient subjects.

JOHN COHEN and C. E. M. HANSEL. The idea of a distribution.

F. W. WARBURTON. The scientific status of mental measurement.

SIDNEY M. JOURARD and PAUL F. SECORD. Body-cathexis and personality.

AGNES CRAWFORD. The Dvorine pseudo-isochromatic plates.

PUBLICATIONS RECENTLY RECEIVED.

OTHER PUBLICATIONS RECEIVED.

EDITORIAL NOTE.

ERRATUM.

Vol. XLVI. Part 3 August 1955 $3.50 net.

B. A. FARRELL. On the limits of experimental psychology.

JOHN COHEN and C. E. M. HANSEL. The idea of independence.

PER SAUGSTAD. Problem-solving as dependent on availability of functions.

K. LOVELL. A study of the problem of intellectual deterioration in adolescents and young adults.

A. W. HEIM. Adaptation to level of difficulty in intelligence testing.

JOSEPH SANDLER. A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*.

CORRESPONDENCE.

EDITORIAL NOTE.

PUBLICATIONS RECENTLY RECEIVED.

The subscription price per volume, payable in advance, is $10 net (post free).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS

Bentley House, Euston Road, London, N. W. 1

ZEITSCHRIFT FÜR
EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

This journal is devoted to experimental investigations covering the entire field of psychology. In addition, up-to-date reviews of current books on psychology are published.

For many years, a strictly scientific journal of psychology has been lacking in German-speaking countries. In taking over this task the "Zeitschrift für experimentelle und angewandte Psychologie" endeavours to continue the tradition of the former, and now extinct, German journals of scientific psychology.

All articles are provided with summaries in English and French.

Herausgeber für Deutschland
Prof. Dr. J. von Allesch, Göttingen; Prof. Dr. Ph. Lersch, München

Herausgeber für Österreich
Prof. Dr. H. Rohracher, Wien

Herausgeber für die Schweiz
Prof. Dr. R. Meili, Bern

Redaktion
Dr. C. J. Hogrefe, Göttingen

Ständige Mitarbeiter

Prof. Dr. W. Arnold, Nürnberg; Prof. Dr. H. Biäsch, Zürich; Prof. Dr. C. Bondy, Hamburg; Prof. Dr. E. Brunswik, Berkeley USA; Prof. Dr. H. Düker, Marburg; Prof. Dr. G. Ekman, Stockholm; Prof. Dr. Th. Erismann, Innsbruck; Prof. Dr. H. J. Eysenck, London; Prof. Dr. R. Heiß, Freiburg i.B.; Prof. Dr. B. Herwig, Braunschweig; Prof. Dr. H. Hetzer, Marburg; Prof. Dr. E. Kretschmer, Tübingen; Prof. Dr. O. Kroh, Berlin; Prof. Dr. W. Metzger, Münster; Prof. Dr. P. Pichot, Paris; Prof. Dr. E. Rausch, Frankfurt; Prof. Dr. G. Révész, Amsterdam; Prof. Dr. E. Sander, Berlin; Prof. Dr. U. Undeutsch, Köln; Prof. Dr. A. Wellek, Mainz; Prof. Dr. K. Wilde, Göttingen.

The journal appears quarterly and contains 160 pages.

Subscription is DM 64.— per volume and DM 18.— for single numbers.

VERLAG FÜR PSYCHOLOGIE · DR. C. J. HOGREFE · GÖTTINGEN
(Postfach 414, Germany)

PSYCHOLOGICAL REVIEW

YEAR	VOLUME	AVAILABLE NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
1894	1	–	2	–	4	5	6	$1.50	——
1895	2	–	–	3	4	5	6	$1.50	——
1896	3	–	–	–	–	–	–	——	——
1897	4	1	–	–	–	–	6	$1.50	——
1898	5	–	–	3	–	5	–	$1.50	——
1899	6	–	–	–	–	–	–	$1.50	——
1900	7	1	–	–	–	–	–	$1.50	——
1901	8	1	2	–	–	–	–	$1.50	——
1902	9	–	2	–	–	–	–	$1.50	——
1903	10	1	2	–	–	–	–	$1.50	——
1904	11	1	–	–	4	& 5	6	$1.50	——
1905	12	1	2	& 3	4	5	–	$1.50	——
1906	13	–	–	3	4	5	6	$1.50	——
1907	14	1	2	–	–	–	–	$1.50	——
1908	15	–	–	–	–	–	–	——	——
1909	16	1	–	3	4	5	–	$1.50	——
1910	17	1	2	3	–	–	6	$1.50	——
1911	18	1	2	3	4	5	–	$1.50	——
1912	19	1	2	3	4	5	6	$1.50	$8.00
1913	20	1	2	3	4	5	6	$1.50	$8.00
1914	21	–	2	3	4	5	6	$1.50	——
1915	22	1	2	3	4	5	6	$1.50	$8.00
1916	23	1	–	–	4	5	–	$1.50	——
1917	24	–	–	–	4	–	–	$1.50	——
1918	25	–	2	3	4	5	6	$1.50	——
1919	26	1	2	3	4	5	6	$1.50	$8.00
1920	27	1	–	–	–	–	–	$1.50	——
1921	28	–	2	–	–	–	6	$1.50	——
1922	29	1	–	–	4	–	–	$1.50	——
1923	30	1	–	3	–	–	–	$1.50	——
1924	31	1	2	3	4	5	6	$1.50	$8.00
1925	32	–	2	3	–	5	–	$1.50	——
1926	33	1	–	3	4	5	6	$1.50	——
1927	34	1	2	3	4	5	6	$1.50	$8.00
1928	35	1	2	3	4	5	6	$1.50	$8.00
1929	36	1	2	3	4	5	6	$1.50	$8.00
1930	37	1	2	3	4	5	6	$1.50	$8.00
1931	38	1	2	3	4	5	6	$1.50	$8.00
1932	39	1	2	3	4	5	6	$1.50	$8.00
1933	40	1	2	3	4	5	6	$1.50	$8.00
1934	41	1	2	3	4	5	6	$1.50	$8.00
1935	42	1	2	3	4	5	6	$1.50	$8.00
1936	43	1	2	3	4	5	6	$1.50	$8.00
1937	44	1	2	3	4	5	6	$1.50	$8.00
1938	45	1	2	3	4	5	6	$1.50	$8.00
1939	46	1	2	3	4	5	6	$1.50	$8.00
1940	47	–	–	3	–	–	–	$1.50	——
1941	48	–	–	–	–	–	6	$1.50	——
1942	49	–	2	3	4	5	6	$1.50	——
1943	50	1	2	3	4	5	6	$1.50	$8.00
1944	51	1	2	3	4	5	6	$1.50	$8.00
1945	52	1	2	3	4	–	6	$1.50	——
1946	53	1	2	3	4	5	6	$1.50	$8.00
1947	54	1	2	–	4	5	6	$1.50	——
1948	55	1	2	3	4	5	6	$1.50	$8.00
1949	56	1	2	3	4	5	6	$1.50	$8.00
1950	57	1	2	3	4	5	6	$1.50	$8.00
1951	58	1	2	3	4	5	6	$1.50	$8.00
1952	59	1	2	3	4	5	6	$1.50	$8.00
1953	60	1	2	3	4	5	6	$1.50	$8.00
1954	61	1	2	3	4	5	6	$1.50	$8.00
1955	62	By subscription $6.50, foreign $7.00						$1.50	——

Postage prepaid on U. S. orders. Add $.50 per volume on foreign orders. All stock subject to prior sale.
The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Current subscriptions and orders for back numbers should be addressed to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N.W. Washington 6, D. C.

Psycholinguistics

A Survey of Theory and Research Problems

Report of the 1953 Summer Seminar Sponsored by the Committee on Linguistics and Psychology of the Social Sciences Research Council

CHARLES E. OSGOOD, *Editor* THOMAS A. SEBEOK, *Associate Editor*

With a Foreword by JOHN W. GARDNER

JOHN B. CARROLL
SUSAN M. ERVIN
JOSEPH H. GREENBERG
JAMES L. JENKINS
FLOYD G. LOUNSBURY
LEONARD D. NEWMARK
SOL SAPORTA
DONALD E. WALKER
KELLOGG WILSON

Price, $2.50

Order from

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street N. W.
Washington 6, D. C.